Hardeweg, Bernd; Wagener, Andreas; Waibel, Hermann

Conference Paper

Towards Comparative and Aggregate Vulnerability: An Analysis of Welfare Distributions in Rural Provinces in Thailand and Vietnam

Proceedings of the German Development Economics Conference, Hannover 2010, No. 14

Towards Comparative and Aggregate Vulnerability: An Analysis of Welfare Distributions in Rural Provinces in Thailand and Vietnam

Bernd Hardeweg*, Andreas Wagener♣, and Hermann Waibel♠

Leibniz University of Hannover

Faculty of Economics and Management

Königsworther Platz 1

30167 Hannover, Germany.

Abstract

Several measures of vulnerability to poverty have been suggested in the literature. In practise, only little is known about the robustness of vulnerability comparisons based on these often quite specific measures. The theory of stochastic orders can be applied to shed some light on such issues.

In the DFG research project "Impact of Shocks on the Vulnerability to Poverty: Consequences for Development of Emerging Southeast Asian Economies" (DFG FOR 756), an extensive panel survey was carried out in six rural provinces of Thailand and Vietnam in 2007. We establish cumulative distribution functions for income and consumption at the provincial level and search for stochastic dominance relations between these distributions. Our comparisons allow for initial, but quite robust conclusions on welfare and provide benchmarks for assessing the vulnerability to poverty in the research regions.

Financial support by the German Research Foundation (DFG) is gratefully acknowledged.

* *Corresponding author*. Institute of Development and Agricultural Economics. E-mail: hardeweg@ifgb.uni-hannover.de.

♣ Institute of Social Policy. E-mail: wagener@sopo.uni-hannover.de.

♠ Institute of Development and Agricultural Economics. E-mail: waibel@ifgb.uni-hannover.de.

1. Introduction

The World Development Report 2000/2001 has raised awareness for the dynamic aspects of poverty in developing countries. Since then, the concept of vulnerability to poverty – as an *ex-ante*, forward-looking approach -- has received much attention in the community of researchers and practitioners of development. In recent years, a number of concepts of vulnerability and associated indicators have been developed without, however, reaching a consensus on the relative merits of these concepts. The most prominent among these concepts are:

- Vulnerability can be understood as *expected poverty* and measured by the probability that a household will be below the poverty line in some future period (e.g. Pritchett *et al.*, 2000). This approach has been underlying most empirical applications to date. [1] It can be generalized by giving higher weight to more severe degrees of poverty, a prominent example being vulnerability indices based on FGT measures of poverty.
- Vulnerability may be regarded as a *low level of expected utility*: a shortfall of an household's expected utility below some threshold level, which is meant to represent a socially acceptable minimum level of (expected) well-being (Ligon and Schechter, 2003).[2]
- Individual vulnerability to poverty as developed by Calvo and Dercon (2005), measured by an index of *expected deprivation*, accounting for both the probabilities of negative future events and their severity.

For practical applications, each of these concepts can be mathematically represented by (a class of) expected values of functions of income or consumption. The concepts of Calvo and Dercon and the expected poverty measures rely on censored incomes. Where

1 See, e.g., Chaudhuri *et al.* (2002) or Christiaensen and Subbarao (2005). Chaudhuri *et al.* (2002) calculate predicted income or consumption from three major components: (i) household (and village) characteristics, (ii) the risks and shocks faced and (iii) the risk coping strategies applied. The prominence of their approach is driven by the availability of data. Often household living standard surveys are used to estimate an income or consumption function interpreting the residuals of the estimates as an expression of idiosyncratic shocks and risks (Chaudhuri *et al.*, 2002), while information on covariate shocks are taken from secondary data like rainfall records (e.g., Dercon, 2004).

2 The Ligon-Schechter approach takes into account the household's degree of risk aversion and allows a decomposition of vulnerability into a risk- and a poverty component.

censoring treats income or consumption levels above certain threshold levels (the *poverty line*) as irrelevant; this reflects the focus axiom of poverty measurement.

Given the multitude of approaches (different concepts with a broad class of measures within each concept, and a continuous range of possible poverty lines), it may appear to the superficial observer that "vulnerability" is a woolly idea, pliable at the discretion of the analyst. Applied studies, thus, appear to invite criticism for the specific concepts of vulnerability or poverty lines they are using.

Such verdicts may be too quick, however. Given that many applied measures share the common mathematical structure of expected values (for precise details, see Section 2), they are in fact quite closely related. These relations can be traced back to the idea of stochastic orderings (surveyed, e.g., in Shaked and Shanthikumar, 2006). The idea of this paper is to utilize the concept of stochastic orderings to compare various income distributions with respect to their vulnerability to poverty. In an application to income and consumption distributions in six rural provinces of Thailand and Vietnam, we show that this idea can indeed be set to work also in practise.

We build on a strand of theoretical literature in welfare economics that aims at rendering comparisons of distributions more ethically robust, by making judgements only when *all* members of a possibly wide class of indices for poverty, inequality, vulnerability, or more generally social welfare lead to the same conclusion, rather than focussing on one particular index. Such robustness appears warranted as a safeguard against lack of agreement on a precise poverty line or on a vulnerability criterion, inequality index or social welfare function. Criticism on specific welfare indices dates back at least to Sen (1976) who criticised the use of the headcount ratio and the poverty gap in poverty measurement as disregarding the intensities and the depth of poverty, respectively. The theory of stochastic orderings provides a unified framework that caters for such concerns. It has been fruitfully and frequently applied in comparative measurement of welfare, inequality, and poverty (for a recent survey, see Duclos and Araar, 2006). Given the rather close relations between vulnerability and social welfare (cf. the understanding of vulnerability as a shortfall in expected utility in Ligon and Schechter, 2003), deprivation and poverty (see Chaudhuri *et al*., 2003, or Calvo and

Dercon, 2005), it appears near at hand to utilize the theory of stochastic orderings also in the vulnerability context.

As a test of its practicality, we apply this idea to data collected in the DFG research project "Impact of Shocks on the Vulnerability to Poverty: Consequences for Development of Emerging Southeast Asian Economies" (DFG FOR 756). This project is among the first to develop a comprehensive empirical data base for the measurement of vulnerability. In 2007 and 2008, a panel survey of some 4400 households was carried out in six rural provinces of Thailand and Vietnam (Hardeweg and Waibel, 2009).[3] An extensive questionnaire was developed that allows us to establish cumulative distribution functions of income and consumption at the level of provinces. Using the 2007 wave, we search for stochastic dominance relations (in the first, second, and third order) between these provincial distributions. Such comparisons allow for initial, but quite robust conclusions on welfare; they provide benchmarks for assessing the vulnerability to poverty of the target population.

Our general insight is that the theory of stochastic orderings provides indeed useful tools for vulnerability analysis. More specifically to the case study, our results are as follows: Provincial distributions for *consumption* within Thailand and Vietnam can, up to certain income thresholds, be ranked by second (or third) degree stochastic dominance criteria, implying that the dominated distributions exhibit, below the threshold incomes, higher degrees of vulnerability for all inequality-averse (respectively, downside inequality-averse) measures. Similarly clear rankings are not available for *income* distributions. Vulnerability assessments, thus, appear to be more robust if made on a consumption rather than on an income basis.

The rest of this paper is structured as follows: Section 2 outlines the notion of stochastic dominance and reviews its connections to vulnerability and related issues. Section 3 reports on the data we use from Vietnam and Thailand and discusses our methodology and its shortcomings. Section 4 presents the dominance comparisons for the sampled provincial income distributions. Section 5 concludes.

---

3 Here we only utilize a small part of the available data. In fact, the full survey includes data on village and household characteristics, income and consumption, shocks experienced in the past, expected negative future events, including the probability of occurrence and their severity, *ex ante* risk coping strategies and the households degree of risk aversion.

2. Stochastic Dominance and Vulnerability

In this section we briefly review the relations between stochastic dominance and (static) measures for poverty and vulnerability. Most of the results on welfare and poverty dominance reported here can be found in Foster and Shorrocks (1988a, 1988b), Davidson and Duclos (2000), or Duclos and Araar (2006, esp. Chapter 10).

2.1 Notation and concepts

We will consider income distributions that have non-negative support and finite (positive) means. We use the term ‘‘income’’ to signify a monetary measure of individual welfare. In our application we also use consumption as an indicator of individual well-being; all concepts introduced here apply to consumption distributions as well (and, in fact, most would *mutatis mutandis* also apply to non-money measures of individual wellbeing). We denote the cumulative distribution function (CDF) of income distribution Y by $F_Y(y)$.[4] Unless stated otherwise, terms "increasing", "decreasing", "convex" etc. are employed in the non-strict sense throughout. Improper integrals and expectations (demoted by E) are assumed to exist whenever they are written. For (non-negative) distributions, it is well-known that $E(Y)=\int_0^\infty (1-F_Y(y))\,\mathrm{d}y$.

a) Stochastic Dominance

Let $D_Y^1(y)=F(y)$ and, for any integer $k>1$, $D_Y^k(y)=\int_0^y D_Y^{k-1}(x)dx$. Given two distributions with CDFs F_A and F_B, distribution B is said to *stochastically dominate* distribution A stochastically at order k if $D_A^k(y)\geq D_B^k(y)$ for all $y\in\mathbb{R}$; for strict dominance, this inequality is required to hold strictly over some interval of positive measure. Suppose that a poverty line was established at some income level $z>0$. Then distribution B is said to stochastically dominate A at order k *up to the poverty line* z if

4 Unless confusion can arise, we will omit the subscript to *F*. For our purposes, there is no need to distinguish between continuous and discrete distributions. In the empirical analysis, of course, a discrete distribution is used.

$D_A^k(y) \geq D_B^k(y)$ for all $y \leq z$; this is sometimes referred to as k-th order *poverty dominance*.

Stochastic dominance properties are nested: k-th order stochastic dominance [up to z] implies, but is not implied by, (k+t)-th order stochastic dominance [up to z] for all positive integers t. Orderings by stochastic dominance are partial; non-comparability between distributions are not uncommon.

b) Social Welfare

For continuous (and sufficiently many times differentiable) function $u : \mathbb{R}_+ \to \mathbb{R}$, consider the class $\mathbf{U}$ of welfare functions of the form

$$U(F) = \int_0^\infty u(y) \mathrm{d}F(y) . \qquad (1)$$

Denote by $\mathbf{U}_k$ ($k = 1, 2, ...$) the subset of $\mathbf{U}$ that is based on functions u that satisfy $(-1)^{s+1} u^{(s)}(y) \geq 0$ for all $s \leq k$ and all y (with $u^{(s)}$ denoting the s-th derivative of u).

c) Poverty

For real-valued x, we shall write $x_+ = \max\{x, 0\}$. If $z > 0$ is a poverty line, then $(z - y)_+$ is called the poverty gap at income level y; using the notion of censored incomes $y^* = \min\{y, z\}$, we can express the poverty gap equivalently as $z - y^*$. Poverty indices are frequently based on poverty gaps. Together with the so-called focus axiom (only incomes at or below the poverty line matter in poverty assessment), poverty indices can typically by expressed as

$$P(z, F) = \int_0^\infty p[(z - y)_+] \mathrm{d}F(y) \qquad (2)$$

where p satisfies $p(0) = 0$. For continuous (and sufficiently many times differentiable) function $p : \mathbb{R} \to \mathbb{R}$ with $p(0) = 0$, consider the class $\mathbf{P}$ of poverty indices of form (2). Denote by $\mathbf{P}_k$ ($k = 1, 2, ...$) the subset of $\mathbf{P}$ that is based on functions p that satisfy $p^{(k)}(x) \geq 0$ for all x, and $p^{(s)}(0) = 0$ for all $s < k$.

d) Vulnerability

A measure of vulnerability to poverty is meant to capture the notion that higher values of that index indicate a larger risk of being in dire straits in the future. A number of measures have been proposed in the literature. A consensus on the right measure has, however, not yet been reached.

Ligon and Schechter (2003) take a utilitarian standpoint and view vulnerability as *low expected utility,* where "low" is defined relative to some minimum socially acceptable utility level. Specifically, their measure of (individual) vulnerability is:

$$V^{LS}(z,F) = u(z) - \int_0^{\infty} u(y)dF(y)\,, \tag{3a}$$

where $u(x)$ is an increasing and concave (utility) function. Variable z is meant to represent a certainty-equivalent income or consumption such that, if an individual had that for certain that consumption level or a higher one, it would not be regarded as vulnerable. This choice of z is analogous to the choice of a "poverty line" in the literature on poverty measurement. For vulnerability comparisons between two distributions, the poverty line does not matter at all in (3a). Moreover, the Ligon-Schechter measures always employ the full support of the distribution; no censoring takes place (which implies a violation of the focus axiom). These crucial differences between the Ligon-Schechter approach to vulnerability and others render comparative vulnerability a far more demanding concept for the former.

Calvo and Dercon (2005, 2007) propose five *axiomatic* properties that an index for individual vulnerability should satisfy; they and show that these axioms are altogether if there exists a decreasing and convex function v such that

$$V^{CD}(z,F) = \int_0^{\infty} v\left(\min\left\{\frac{y}{z},1\right\}\right)dF(y)\,. \tag{3b}$$

Specifically, Calvo and Dercon (2007) suggest to set $v(x) = 1 - x^{\beta}$ for some $\beta \in [0,1]$ (this specification will, however, not play a big role here).

Quite frequently, vulnerability is understood as *expected poverty* (Ravallion, 1998; Chaudhuri, 2003). Most of the applied literature (see, e.g., Chaudhuri et al., 2000, or Christiaensen and Subbarao, 2005) employs the additive FGT poverty indices (due to Foster et al., 1984). They are a special case of (1) and define vulnerability by

$$V^{EP}(z,F;\alpha) = z^{1-\alpha}\int_0^\infty (z-y)_+^{\alpha-1}\,dF(y), \tag{3c}$$

where $\alpha \geq 1$ measures the relative weights attached to individuals hit more severely by poverty. Since $D^k(y) = \frac{1}{(k-1)!}\int_0^y (y-x)^{k-1}\,dF(x)$, one gets that

$$V^{EP}(z,F;\alpha) = z^{1-\alpha} D_F^{\alpha}(z)$$

for integer values of α. This representation allows very easy connections of expected poverty in the FGT sense and stochastic dominance (see below).

We call a vulnerability index *focused* if it only depends on those parts of a distribution that are below the poverty line. Precisely, V is said to be focused if, for all $z > 0$, and all distributions F_A, F_B, if $F_A(y) = F_B(y)$ for all $y \leq z$, we also have $V(z,F_A) = V(z,F_B)$. The Calvo-Dercon measures (3b) and expected poverty indices (3c) and (3d) are focused; the Ligon-Schechter measures (3a) are not focused.

2.2 Connections

The following items illustrate the relationships between stochastic dominance and some concepts used in the literature on poverty and vulnerability measurement; they are all well-known from the literature:

a) First-order stochastic dominance (FSD)

If distribution A is first-order stochastically dominated by distribution B up to poverty line *z*, then the headcount ratio (i.e., the proportion of individuals below the poverty line) is always greater in A than in B: $F_A(x) \le F_B(x)$ for all $x \le z$. Moreover, if distribution B first-order stochastically dominates distribution A, then the headcount ratio is, at *any* poverty level, lower in B then in A.

If distribution B first-order stochastically dominates distribution A, all welfare comparisons that are based on increasing functions of income identify B as the preferable distribution: $U(F_B) \ge U(F_A)$ for all $U \in \mathbf{U}_1$. More generally, if distribution B first-order stochastically dominates distribution A up to poverty line z, then, for all increasing functions $\phi : \mathbb{R}_+ \to \mathbb{R}$,

$$E_B(\phi(y) | y \le z) \ge E_A(\phi(y) | y \le z).$$

In addition, at any poverty line z, all poverty comparisons that are based on functions that are increasing in the poverty gap exhibit lower poverty levels in B then in A: $P(z, F_B) \le P(z, F_A)$ for all $P \in \mathbf{P}_1$ and all $z > 0$.

The status of these results becomes clearer if we consider the absence of FSD (analogous interpretations will apply to the connections for higher orders of dominance reported below): If neither of two distributions A and B first-order stochastically dominates the other [up to some poverty level z], then there exist welfare indices $U^1, U^2 \in \mathbf{U}_1$ such that $U^1(F_B) \ge U^1(F_A)$, but $U^2(F_A) \ge U^2(F_B)$. Likewise, if neither of the distributions A and B first-order stochastically dominates the other [up to some poverty level z], then there exist poverty indices $P^1, P^2 \in \mathbf{P}_1$ such that $P^1(z, F_B) \le P^1(z, F_A)$ but $P^2(z, F_A) \le P^2(z, F_B)$. Moreover, it may occur that $P(z^1, F_B) \le P(z^1, F_A)$ but $P(z^2, F_A) \le P(z^2, F_B)$ for a given index $P \in \mathbf{P}_1$ at two distinct poverty lines $z^1, z^2 < z$. Hence, without first-order stochastic dominance, welfare or poverty comparisons between distributions vary with the specific valuation functions (or poverty lines) chosen by the investigator within the classes $\mathbf{U}_1$ or $\mathbf{P}_1$. For sure, classes $\mathbf{U}_1$ or $\mathbf{P}_1$ are very wide (and potentially unreasonably so). Hence measure-independence of vulnerability or poverty comparisons within these large classes is a highly demanding (and potentially unreasonable) property that empirically can only be expected to hold in very few instances, as we shall also see below in our data from Vietnam and Thailand.

For first-order stochastic dominance, the following results obtain for vulnerability measures: If distribution B (fully) first-order stochastically dominates distribution A, then, at all poverty lines, then for all z,

$$V^{LS}(z,F_A) \geq V^{LS}(z,F_B)\,, \tag{4a}$$

$$V^{CD}(z,F_A) \geq V^{CD}(z,F_B)\,, \tag{4b}$$

$$V^{EP}(z,F_A;\alpha) \geq V^{EP}(z,F_B;\alpha) \text{ for all } \alpha \geq 1\,. \tag{4c}$$

It should be noted that for Ligon-Schechter indices (3a), Calvo-Dercon indices (3b) and expected poverty measures (3c) with $\alpha > 1$ first-order stochastic dominance of A by B is not a necessary, but only a sufficient requirement for the comparisons above to hold (this is due to the facts that the LS-index uses a concave utility function while the CD-index and the expected poverty indices with $\alpha > 1$ uses a convex measure for the severity of poverty; see below).

If distribution B first-order stochastically dominates distribution A up to income level $\bar{z}$ (i.e., if B poverty-dominates A in the first order up to a poverty level $\bar{z}$), then (4b) and (4c) hold for all poverty lines $z \leq \bar{z}$. This is not true, however, for (4a) since the Ligon-Schechter indices lack focus.

b) Second-order stochastic dominance (SSD)

If distribution *A* is second-order stochastically dominated by distribution *B* [up to poverty line *z*], then the aggregate poverty gap [for all poverty levels at or below z] is always greater in A than in B: $D_A^2(x) \geq D_B^2(x)$ for all x [for all $x \leq z$].

If distribution B second-order stochastically dominates distribution A, then all welfare comparisons that are based on increasing and concave functions of income identify B as the preferable distribution: $U(F_B) \geq U(F_A)$ for all $U \in \mathbf{U}_2$. More generally, if distribution B second-order stochastically dominates distribution A up to poverty line z, then, for all increasing and concave functions $\phi:\mathbb{R}_+ \to \mathbb{R}$,

$$E_B(\phi(y)\,|\,y \le z) \ge E_A(\phi(y)\,|\,y \le z).$$

If distribution B stochastically dominates distribution A in the second order, then, at all poverty lines z, all poverty comparisons that are based on functions that are increasing and convex in the poverty gap exhibit lower poverty levels in B then in A: $P(z,F_B) \le P(z,F_A)$ for all $P \in \mathbf{P}_2$ and all $z > 0$.

As most common measures for vulnerability attach higher weight to larger shortfalls below the poverty line (i.e., they are based on convex poverty measures), almost all results for vulnerability extend from first-order stochastic dominance to second order stochastic dominance. Specifically, if distribution B second-order stochastically dominates distribution A, then for all z,

$$V^{LS}(z,F_A) \ge V^{LS}(z,F_B); \tag{5a}$$

$$V^{CD}(z,F_A) \ge V^{CD}(z,F_B); \tag{5b}$$

$$V^{EP}(z,F_A;\alpha) \ge V^{EP}(z,F_B;\alpha) \text{ for all } \alpha \ge 2. \tag{5c}$$

Different from FSD, dominance in the second order is also a necessary condition for the comparisons to hold. Similar to FSD, if we only have second-order poverty dominance up to some level $\bar{z}$, then (5b) and (5c) hold for all poverty lines $z \le \bar{z}$, while (5a) does not.

Defining the Generalized Lorenz Curve for distribution F as

$$GL(q;F) = \int_0^q F^{-1}(t)\mathrm{d}t$$

(for $0 \le q \le 1$) one gets that, if distribution B second-order stochastically dominates distribution A, then the Generalized Lorenz Curve of B is never below that of A:

$$GL(q;F_B) \ge GL(q;F_A)$$

for all q. As comparisons of Generalized Lorenz Curves boil down to comparisons of "normal" Lorenz curves if distributions have a common mean, we get that for distributions with equal means, second-order stochastic dominance is equivalent to unambiguously lower inequality in the Lorenz sense.

c) General

Many of the results reported above easily generalize to any desired stochastic order k. E.g., the following statements are equivalent:

- Distribution B k-th stochastically dominates distribution A [up to poverty level $\bar{z}$];
- $P(z,F_B) \leq P(z,F_A)$ for all $P \in \mathbf{P}_k$ for all $z > 0$ [respectively, all $z \leq \bar{z}$];
- $V^{EP}(z,F_B;\alpha) \leq V^{EP}(z,F_A;\alpha)$ for all $\alpha \geq k$ for all $z > 0$ [respectively, all $z \leq \bar{z}$];
- $U(F_B) \geq U(F_A)$ for all $U \in \mathbf{U}_k$ [respectively, $\int_0^z u(x)dF_B(x) \geq \int_0^z u(x)dF_A(x)$ for all $U \in \mathbf{U}_k$ and all $z \leq \bar{z}$].

In our application below, we confine ourselves to maximally third-order stochastic dominance (TSD; $k = 3$). This order is quite of some relevance for poverty assessments; social welfare functions $U \in \mathbf{U}_3$ are called downside risk (or inequality) averse.

2.3 Summary

A stochastic dominance test of two income distributions, A and B, has three possible outcomes: A is dominated by B; B is dominated by A; or neither holds. Applied to the vulnerability-to-poverty context, such assessments can be translated into vulnerability rankings: For all vulnerability measures within a certain class (depending on the stochastic order), B is preferred to A; A is preferred to B; or the test cannot discriminate between the two distributions. Recalling that dominance properties are nested, when the degree of stochastic dominance is increased, the frequency of the non-comparability outcome is reduced. This comes at the cost, however, that the preference criterion gets weaker and the class of vulnerability measures more restricted.

With a lower degree of stochastic dominance between two distributions, more general (or less arbitrary) comparisons can be made about their relative vulnerability; a lower degree of stochastic dominance implies a weaker dependence of vulnerability comparisons on the specific measures used.

The different orders of dominance correspond to increasing restrictions on the shape of the social welfare function and/or the degree of "aversion" that vulnerability concepts assign to increasingly severe levels of poverty. However, these restrictions are non-parametric; in particular, they do not presuppose any specific functional forms in measures like (1), (2), and (3a) through (3c).

Seen from a stochastic orderings perspective, vulnerability indices of types (3b) and (3c) (for $\alpha > 1$) appear quite similar: if two distributions can be ranked with respect to SSD [up to a certain poverty line], then all of them would agree in their vulnerability ranking of these distributions. Without SSD, they might lead to divergent comparisons, however.

3. Data and Method

3.1 Data description

In the context of the DFG research project "Impact of Shocks on the Vulnerability to Poverty: Consequences for Development of Emerging Southeast Asian Economies" (DFG FOR 756), in 2007 and 2008 a panel survey of some 4400 rural households was carried out in six provinces of Thailand and Vietnam. These provinces are Buriram (BR), Nakhon Phanom (NP), and Ubon Ratchathani (UR) in Thailand, and Dak Lak (DL), Ha Tinh (HT), and Thua Thien Hue (TH) in Vietnam. Provinces were selected purposively to include peripheral areas in the poorest region of Thailand and provinces representing different levels of economic development of the Central Coast and Central Highland regions of Vietnam.

Households were selected in a three-stage random sampling procedure. In the first stage, sub-districts were chosen with a probability proportional to size from strata, defined by the provinces in Thailand. In Vietnam, the stratification into three agro-ecological zones within provinces with disproportional sample allocation was applied in order to ensure sufficient sample size in the less densely populated highland areas. In the second stage, two villages were sampled from each sub-district with probability proportional to size, before a fixed sample of ten households was selected from each village cluster in the ultimate stage. As a result we obtain a sample that is representative for the rural population of the selected provinces.

An extensive questionnaire generated data that allows us to establish, for two consecutive waves, cumulative distribution functions of income and consumption at the level of provinces. In this paper we search for stochastic dominance relations between these distributions for the 2007 wave. Such comparisons allow for initial, but quite robust conclusions on welfare; they provide benchmarks for assessing the vulnerability to poverty of the target population.

3.2 Method

In this paper we use raw data for sampled income and consumption distributions over individuals of six rural areas in Thailand and Vietnam, considering sampling weights to account for unequal sampling probabilities (see Section 3.1). We take these empirical distributions and subject them to dominance comparisons. We then interpret our observations in terms of comparative vulnerability (see Section 4 below).

The main purposes of this paper are (i) to present summaries for the income distributions in the sample regions and (ii) to show how such summary information can, *in principle*, be used for vulnerability comparisons. Given this limited scope of our paper, we should be frank about the limitations of our method and analysis. Two major concerns are:[5]

a) Measures of vulnerability: Indices such as (3a) through (3c) are thought to measure *individual* vulnerability to poverty. Moreover, vulnerability is very much an *ex ante* concept. As our data, we use sampled empirical (i.e., *ex post*) distributions of consumption or income over quite large groups of households as the basis for vulnerability comparisons. At best, we might interpret such an analysis as the construction of a hypothetical, representative inhabitant for each province who uses an observed income distribution for the population in his province as a predictor for his own future, personal income distribution. Such a veil-of-ignorance approach is, of course, questionable (but, on practical grounds, unavoidable).

b) Econometric issues: We base comparisons among distributions on (more or less) unpolished sample data. We do not include any econometric analysis. The literature has established various statistical tests for stochastic dominance (see, e.g., Davidson and Duclos, 2000), both with independent as with independent distributions. While the empirical distribution function is a good estimator for the (unknown) population cumulative distribution function (Anderson, 1996; Davidson and Duclos, 2000), future versions of this paper should be based on thorough statistical testing of dominance relations between the sampled distributions.

Having noted these caveats, we proceed under the proviso that (i) the sampled distributions and their dominance relations correctly reflect population distributions and the dominance relations among them and that (ii) these population distributions are sufficient statistics for the ex-ante (stochastic) distributions for personal incomes which representative inhabitants in the corresponding province face.

[5] Minor quibbles can also be added: Some households in our sample report negative incomes, which are (at least strictly speaking) not covered in the formal description of Section 2. Moreover, we treat distributions as independent, ignoring that income levels within (and possibly also across) countries are subject to covariate shocks.

4. Results

We obtained our results with the help of the DASD software package for STATA developed by Araar and Duclos (2009). This package is able to account for our rather complex sample design. We report comparative results with respect to first, second and third degree of stochastic dominance between the 2007 distributions of per capita and per adult equivalent daily income and daily consumption. Our results are summarized in Table 1.

Table 1 goes here.

a) First-order stochastic dominance

For a first impression, let us, however, consider Figures 1 and 2 that depict the cumulative distributions of per-capita consumption, separately for the Thai and the Vietnamese provinces.

Figures 1 and 2 go here.

In these graphs, each pair of curves has at least one point of intersection. Hence, the provincial income distributions do not exhibit full first-order dominance relations within the two countries. This implies that comparisons of social welfare or poverty between these provinces are measure- and poverty-line dependent if measures are allowed to come from the (very large) classes of valuation functions $\mathbf{U}_1$ and $\mathbf{P}_1$, respectively (which encompass all welfare functions that are merely increasing in incomes or, respectively, poverty measures that increase with the frequency of poverty occurrences).

To illustrate this, consider the FGT poverty index for $\alpha = 1$ (see (3c)), which is used as a vulnerability index, e.g., in Chaudhuri *et al*. (2002). At poverty line *z*, it indicates a higher vulnerability in province A than in province B if $F_A(z) > F_B(z)$. At any intersection of $F_A(z)$ and $F_B(z)$ this relation changes, however, rendering the vulnerability comparison dependent on the specific choice of *z*. This is the case for the sampled provinces in Thailand and Vietnam.

That we do not find full FSD relations between the provincial distributions is an expected result, given the strength of FSD. Still, for *focused* vulnerability that rely on censored values for income and consumption we can be much more specific in partial vulnerability comparisons by identifying the first point of intersection of pairs of cumulative distribution functions. These earliest crossing points, together with the direction of the crossing, are reported in the FSD-rows of Table 1 for the various categories of income and consumption. E.g., the pair $(<, 3.86)$ reported in the FSD comparison for income per capita between the Thai provinces of Buriram and Nakhon Phanom indicates that the attending distribution for Nakhon Phanom first-order dominates that of Buriram up to the income level of 3.86 $ PPP(2005) (i.e., the cumulative distribution function of incomes in Buriram first crosses that of Nakhon Phanom from below at 3.86 $). Hence, for all poverty lines below 3.86 $ *all* (monotonic) poverty and all focused vulnerability indices would indicate that (a representative individual in) Buriram is poorer and more likely to be vulnerable than (in) Nakhon Phanom. This, of course, encompasses the focused vulnerability measures (3b) and (3c), but excludes the unfocused Ligon-Schechter class (3a).

Given that 3.86 $ is well above all commonly used poverty lines, it appears justified[6] to interpret people in Buriram to be comparatively more vulnerable to poverty *at standard poverty lines* than in Nakhon Phanom, independently of how vulnerability is specifically measured (as long as the measure is focused). Similarly robust claims can also be made for comparisons for consumption vulnerability (both in per capita and per adult terms) between Ha Tinh and Thua Thien Hue as well as between Dak Lank and Thua Thien Hue (with unambiguously less vulnerability in Thua Tien Hue, for all measures and commonly used poverty lines).

A look across the FSD-rows for income in Table 1 indicates, however, that such far-reaching comparisons cannot be made for other pairs of provinces. For income distributions, the first crossing points are all negative and therefore below all commonly used poverty lines. For consumption, the first-crossing points lie around the 1 $- (or 1.25 $-) threshold. This suggests, e.g., that some care should be taken when using the

[6] All caveats mentioned in Section 3.2 still apply.

poverty gap (i.e., the FGT-index with $\alpha = 1$) with these standard poverty lines in measuring vulnerability.[7]

b) Second-order stochastic dominance

For vulnerability comparisons based on the commonly used measures (3b) and (3c), the property of FSD is only partly relevant: If (which is not the case in our sample of Vietnamese and Thai provinces) clear FSD relations prevailed, then the dominated province would always exhibit a higher degree of vulnerability, whatever the measure or poverty line chosen in (3b) and (3c). The non-existence of any such FSD relation, however, does not yet tell us anything about measures (3b) and (3c) since they impose more restrictions than just $\mathbf{U}_1$ and $\mathbf{P}_1$. For that, we need (at least) second-order considerations, which we address now.

For a visual impression, have a look at Figure 3 which presents the integrals over cumulative distribution for per capita consumption in the three Vietnamese provinces under study (the picture for per-adult consumption looks very similar).

Figure 3 goes here.

Interestingly, we see that the consumption distribution of Thua Thien Hue second-order stochastically dominates that of Ha Tinh. Hence, all vulnerability measures of types (3a), (3b), and (3c) (for $\alpha > 1$) would, at any poverty line, agree that an individual is less vulnerable with respect to consumption in Thua Thien Hua than in Ha Tinh. Within the class of Ligon-Schechter measures (with concave *u*), Calvo-Dercon measures (with convex *v*), and FGT measures (with $\alpha > 1$) this comparison is robust for every given poverty line. In addition, the comparison remains, for any poverty line, robust for whatever measure of vulnerability one can conceive of that attaches more severity to low incomes or consumption levels that are farther away from the poverty line.

7 As an illustrative example, take the case of Buriram vs. Nakhon Phanom for consumption per adult and take the first crossing point at 1.18 $ literally. The FGT poverty gap with a poverty line of 1.17 $ would show Nakhon Phanom as the less vulnerable province, a poverty line of 1.19 $ Buriram.

For all pairwise comparisons without full stochastic dominance in the second order, Table 1 again reports the supremal levels of income and consumption up to which poverty dominance in the second order holds. Clearly, these threshold levels are higher than those for FSD, reflecting that second-order comparisons are more robust (while also less general) than FSD ones.

The interpretation of the threshold levels in the SSD case is analogous to the FSD case: E.g., the pair $(<, 12.96)$ reported in the SSD comparison for income per capita between the Thai provinces of Buriram and Nakhon Phanom in Table 1 indicates that the attending distribution for Nakhon Phanom second-order dominates that of Buriram for all poverty lines not higher than 12.96 $ PPP(2005). Below that threshold, all monotonic and convex poverty and focused vulnerability indices would indicate that (a representative individual in) Buriram is more vulnerable than (in) Nakhon Phanom. This again encompasses the classes (3b) and (3c) of focused vulnerability measures (and again excludes the unfocused Ligon-Schechter measure).

Recalling from the FSD case that the threshold poverty line between Buriram and Nakhon Phanom was merely 3.86 $, we see that the vulnerability comparison can now be extended over a far wider range of poverty lines – however at the expense of a more severely restricted class of (focused) vulnerability measures.

As evidenced by the low threshold levels for SSD reported in Table 1, comparisons of income vulnerability are, within the classes of commonly used vulnerability indices, highly sensitive with respect to the measures and poverty lines employed. As with FSD, however, we get that inter-provincial vulnerability comparisons can be established in a more robust way for consumption distributions than for income distributions; the threshold levels are significantly higher and exceed standard poverty lines.

c) Higher-order stochastic dominance

For rankings of income and consumption distributions in Vietnam and Thailand, going beyond second-degree stochastic dominance does not add much insight. Clearly, where we had stochastic dominance in the second order, we also have it in the third-order. Basically, the only interesting new case is the income comparison between Buriram and Ubon, which exhibits stochastic dominance in the third order. It implies, e.g., that *all*

FGT indices (3c) with $\alpha > 2$ would, for any poverty line, identify Ubon as the more vulnerable province than Buriram. The same would hold for all Ligon-Schechter measures (3a) with $u'''(x) > 0$ (i.e., with downside risk aversion; cf. Menezes et al., 1980) and for all Calvo-Dercon indices (3b) with $v'''(x) < 0$ (this includes, e.g., functions $v(x) = 1 - x^{\beta}$ with $\beta \in (0,1)$, as discussed in Calvo and Dercon, 2007).

Moreover, the first crossing points listed in the TSD-rows of Table 1 indicate that most inter-provincial vulnerability and poverty comparisons for all the measures just mentioned would be robust at common poverty lines also for income distributions (and not only, as with SSD comparisons, for consumption distributions).

d) Cross-country and aggregate comparisons

It is tempting (albeit problematic) to rank distributions also for provinces in *different* countries. The final two columns in Table 1 report two such comparisons. The first comparison between Ubon Ratchathani and Dak Lak was selected as it exhibits a rare case of full-range first-order stochastic dominance in per-capita consumption. Hence, *all* measures of vulnerability based on increasing utility functions or decreasing poverty functions would equivocally agree that Dak Lak is worse off. The second comparison is between the two provinces that have the lowest average income in the two countries under study: Nakhon Phanom and Ha Tinh. While no clear-cut comparisons of income distributions emerge and the FSD test for consumption also fails, Nakhon Phanom's dominance over Ha Tinh in the second order for consumption indicates that all commonly used vulnerability indices (including the unfocused Ligon-Schechter measures) would agree that Ha Tinh is the more vulnerable province.

One could also aggregate distributions: Comparing country-wise pooled data, shows that for per-capita consumption, measured in purchasing power parity adjusted US$, the sampled distribution for the Thai provinces stochastically dominates that for the Vietnamese provinces in the first order. This suggests that rural households in Vietnam are more vulnerable than in Thailand according to *all* commonly used indicators of vulnerability and poverty. However, such a conclusion should be taken with utmost caution; it is based on a very small sample only and the methodological issues raised in Section 3.2 are compounded in the aggregation.

It would be a more promising extension to compare income and consumption distributions (at the provincial levels). This could shed light on the important issues such as consumption smoothing, the role of remittances, etc.

e) Taking stock

Our application to Thailand and Vietnam gives rise to the following observations:

- As expected, (meaningful) FSD relations do not emerge from the data. Hence, all comparisons of vulnerability are, at least to some degree, dependent on the measure used and/or on the poverty line employed.
- For the distribution of *consumption*, poverty dominance relations in the second order are not uncommon at (and also above) standard poverty lines. Hence, the focused vulnerability measures proposed by Calvo and Dercon (2005) and all expected poverty measures based in FGT indices would agree on comparative vulnerability assessments for the provinces in our samples at standard poverty lines (but not elsewhere).
- Subject to the restrictions mentioned in the previous item, for the three Thai provinces in the sample the *consumption* vulnerability ranking would be, in order of increasing vulnerability: Nakhon Phanom, Buriram, Ubon Ratchathani. For Vietnam, we obtain: Ha Tinh, Thua Thien Hue, Dak Lak.
- For *incomes*, comparable (weak) dominance relations as for consumption do not seem to exist. As a consequence, conflicting assessments of comparative vulnerability are more likely to occur for income than for consumption.
- Vulnerability rankings may differ when done for income rather than for consumption, as exemplified by the reversal of “<” and “>” in the case Ha Tinh vs. Thua Thien Hue. Hence, the indicator of well-being underlying vulnerability measurement ought to be carefully discussed.
- For the vulnerability comparisons in this paper it does not seem to matter whether we use income and consumption in per-capita or in per-adult terms (all

comparisons are identical). Clearly, this is not a general result as household sizes and compositions can be expected to affect vulnerability.

5. Conclusions

With an application to a large household sample from Thailand and Vietnam, we study whether and to what extent the vulnerability of different target populations can be compared independently of specific definitions of vulnerability indices and poverty lines. We exploit the fact that dominance relations for stochastic orderings (of various types) are closely related to the comparability of income distributions for large classes of measures of vulnerability and poverty.

Our paper is a first (and necessarily imperfect and incomplete) attempt to use ideas from the theory of stochastic orderings to rank income distributions with respect to their degree of vulnerability to poverty. The application to questionnaire-based income and consumption data retrieved from the DFG research project "Impact of Shocks on the Vulnerability to Poverty: Consequences for Development of Emerging Southeast Asian Economies" (DFG FOR 756) in Thailand and Vietnam demonstrates that our idea could in fact be fruitful. The results reported here could serve as benchmarks for future studies on comparative vulnerability. Moreover, the techniques illustrated here could be also applied for different aggregates than provinces (e.g., farm households vs. non-farm households; rice farmers vs. producers of other crops etc.).

Knowledge whether and how income distributions compare to one another with respect to (their degree of) stochastic dominance is tantamount to knowing which (if any) classes of vulnerability measures and concepts would come up with equivocal assessments of the relative vulnerabilities in these distributions. Put a bit more bluntly, it tells us to what extent vulnerability comparisons are driven by the analyst's choice of measure or really "come from the data".

Knowledge about stochastic dominance properties between income distributions is relevant also from a policy perspective. It not only shows where the poor and rich provinces lie but also indicates how politics could intervene should it wish to alter the relative standings of provinces.

E.g., since one distribution first-order stochastically dominating another means that in the latter incomes are lower at all percentiles of the distribution, FSD is indicative of a generally worse economic structure. Then, promoting growth (only) in the province with the dominated distribution would tend to equalize welfare levels (if that were desired). With second-order dominance at common means, the dominating distribution is more unequal (in the Lorenz sense) than the dominated one. Hence, an equalization of welfare levels across provinces requires intra-provincial redistribution from rich to poor in the more unequal province. Likewise, if -- at equal first and second moments – a distribution is dominated by another in the third degree, than is entails a higher degree of skewness or downside risk. The corresponding policy tool would have to be transfer-sensitive, i.e., must obey the principle of diminishing transfers (Kolm, 1976; Davies and Hoy, 1995), stating that a transfer of income or consumption from a richer to a poorer person is considered to be more equalizing the lower it occurs in the income or consumption distribution. Much research on the policy implications of observations of different degrees of vulnerability has to be done.

References

Anderson, G. J., 1996, Nonparametric Tests of Stochastic Dominance in Income Distributions, *Econometrica* 64, 1183-1193.

Araar, Abdelkrim and Jean-Yves Duclos, 2009, “DASP: Distributive Analysis Stata Package”, Version 2.0, PEP, CIRPÉE and World Bank, Université Laval.

Calvo, C., and S. Dercon, 2007, “Vulnerability to Poverty”. CSAE WPS/2007-03 , Centre for the Study of African Economies, Oxford.

Calvo, C., and S. Dercon, 2005, “Measuring Individual Vulnerability”. Discussion Paper No. 229, Department of Economics, University of Oxford.

Chaudhuri, S., 2003, Assessing Vulnerability to Poverty: Concepts, Empirical Methods and Illustrative Examples. Mimeo, Columbia University.

Christiaensen, L.J., and K. Subbarao, 2005, Towards an Understanding of Household Vulnerability in Rural Kenya. *Journal of African Economies* 14, 520-558.

Davidson, R., and J.-Y. Duclos, 2000, Statistical Inference for Stochatic Dominance and for the Measurement of Poverty and Inequality. Econometrica 68, 1435-1464.

Davies, J. B., and M. Hoy, 1995, Making Inequality Comparisons when Lorenz Curves Intersect. *American Economic Review* 85, 980-986.

Dercon, S., 2004, *Insurance against Poverty.* Oxford University Press, Oxford.

Duclos, J.-Y., and A. Araar, 2006, *Poverty and Equity. Measurement, Policy and Estimation with DAD*. Springer, New York.

Foster, J. E., J. Greer, and E. Thorbecke, 1984, “A Class of Decomposable Poverty Measures”, *Econometrica* 52, 761-766.

Foster, J. E., and A. F. Shorrocks, 1988a, ‘‘Poverty Orderings,’’ *Econometrica* 56, 173-177.

Foster, J. E., and A. F. Shorrocks, 1988b, ‘‘Poverty Orderings and Welfare Dominance’’, *Social Choice Welfare* 5, 179-198.

Hardeweg, B., and H. Waibel, 2009, “Collecting Data to Measure Vulnerability to Poverty: An Overview”. Paper presented at the Workshop of the DFG Research Group on Vulnerability to Poverty in Southeast-Asia (Hue, Vietnam). Mimeo, University of Hannover.

Kolm, S.Ch., 1976, “Unequal inequalities I”, *Journal of Economic Theory 12,* 416-442.

Ligon, E., and L. Schechter, 2003, “Measuring Vulnerability”, *Economic Journal* 113, C95-C102.

Menezes, C., C. Greiss, and J. Tressler, 1980, “Increasing Downside Risk”, *American Economic Review* 70, 921-932.

Pritchett, L., A. Suryahadi, and S. Sumarto, 2000, Quantifying Vulnerability to Poverty: A Proposed Measure, with Application to Indonesia. SMERU Working Paper (http://www.smeru.or.id/report/workpaper/vulnerability/vulnerability.htm).

Ravallion, M., 1988, “Expected Poverty under Risk-Induced Welfare Variability”, *Economic Journal* 98, 1171-1182.

Sen, A., 1976, ‘‘Poverty: An Ordinal Approach to Measurement,’’ *Econometrica* 44, 219-231.

Shaked, M., and J. G. Shanthikumar, 2006, *Stochastic Orderings*. Springer, New York.

Shorrocks, A. F., and J. E. Foster, 1987, “Transfer Sensitive Inequality Measures,’’ *Review of Economic Studies* 54, 485-497.

Tables

Table 1: Dominance relations for provincial distributions

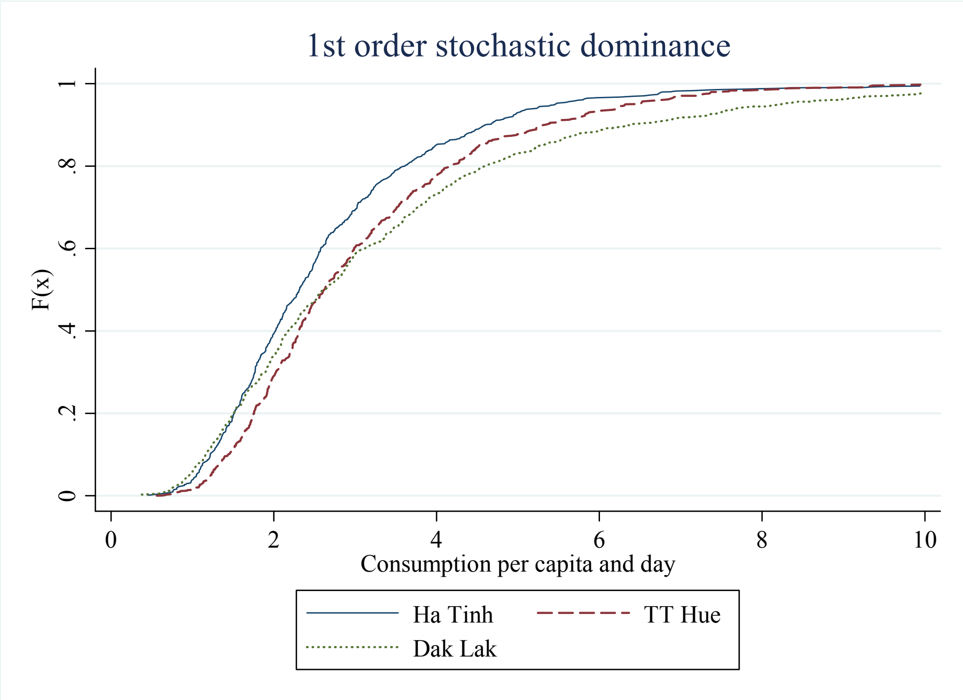
	Thailand			**Vietnam**			**Cross-country (selected)**	
	Buriram - Ubon'	Buriram - Nakhon Phanom	Ubon' - Nakhon Phanom	Ha Tinh - Thua Thien Hue	Ha Tinh - Dak Lak	Thua Thien Hue - Dak Lak	Ubon' - Dak Lak	Nakhon Phanom - Ha Tinh
Income per capita								
FSD	$>$ -7.69	$<$ 3.86	$<$ -5.85	$>$ -0.41	$>$ -1.45	$>$ -2.97	$<$ -0.40	$<$ -1.19
SSD	$>$ -5.43	$<$ 12.96	$<$ 2.52	$>$ 0.28	$>$ 0.06	$>$ -2.05	$<$ 0.67	$<$ 0.21
TSD	$\gg$	$<$ 27.19	$<$ 4.79	$>$ 0.95	$>$ 0.71	$\gg$	$<$ 2.46	$<$ 1.16
Income per adult equivalent*								
FSD	$>$ -10.05	$<$ 5.77	$<$ -8.61	$>$ -1.22	$>$ -1.85	$>$ -3.99	$<$ -0.59	$<$ -1.57
SSD	$>$ -6.76	$<$ 17.34	$<$ 3.51	$>$ 0.41	$>$ 0.10	$>$ -3.08	$<$ 0.97	$<$ 0.30
TSD	$\gg$	$<$ 37.12	$<$ 6.67	$>$ 1.38	$>$ 1.00	$\gg$	$<$ 3.58	$<$ 1.65
Consumption per capita								
FSD	$>$ 0.83	$<$ 0.73	$<$ 0.74	$<$ 9.18	$>$ 1.53	$>$ 2.51	$\gg$	$>$ 0.74
SSD	$>$ 1.72	$<$ 0.75	$<$ 1.90	$\ll$	$>$ 1.97	$>$ 5.11	$\gg$	$\gg$
TSD	$>$ 2.34	$<$ 0.78	$<$ 2.29	$\ll$	$>$ 2.51	$>$ 8.32	$\gg$	$\gg$
Consumption per adult equivalent*								
FSD	$>$ 1.03	$<$ 1.18	$<$ 1.18	$<$ 11.51	$>$ 0.92	$>$ 3.69	**	$>$ 0.68
SSD	$>$ 2.58	$<$ 2.86	$<$ 2.86	$\ll$	$>$ 2.55	$>$ 6.80	**	$\gg$
TSD	$>$ 3.59	$<$ 3.59	$<$ 3.59	$\ll$	$>$ 3.01	$>$ 10.83	**	$\gg$

Notes:
Income and consumption per day in $ PPP(2005).
$A > B$ [$A < B$] denotes dominance of distribution A over B [B over A] in the respective stochastic order up to the value reported in the row below.
$A \gg B$ and $A \ll B$ denote stochastic dominance over the whole range.
* Based on the OECD adult equivalence scale. ** Comparison between Dak Lak and Ubon Ratchathani requires further investigation due to data inconsistencies

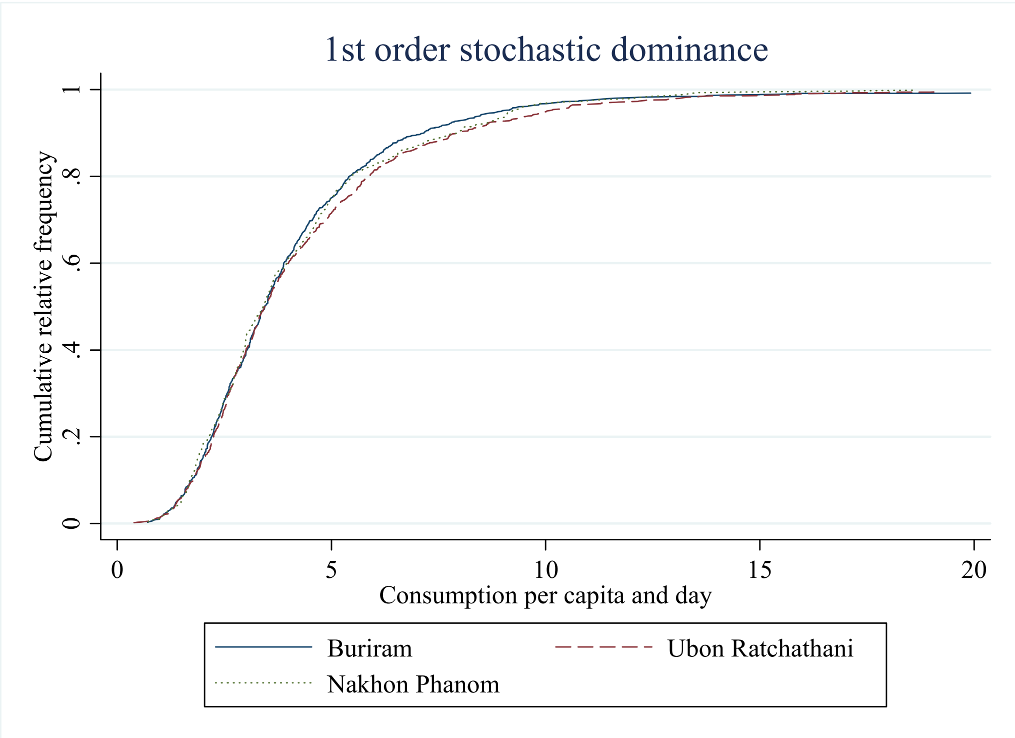
Figures

Figure 1.

A. Distribution of per capita consumption in Vietnamese provinces



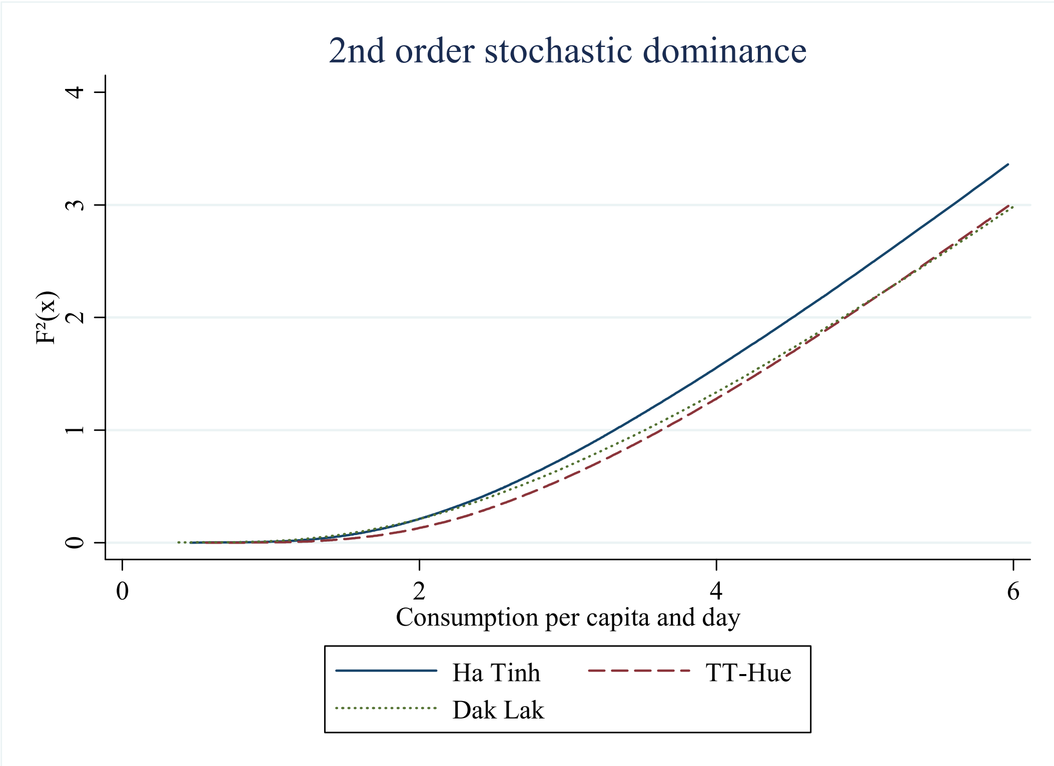
B. Distribution of per capita consumption in Thai provinces



Note: Consumption per capita and day in $ PPP(2005).

Figure 2. SSD for consumption per capita in Vietnam

A. Lower part of the distribution:



B. Full range of the distribution: