

Households' Vulnerability to Covariate and Idiosyncratic Shocks

Isabel Günther* and Kenneth Harttgen†

April 11, 2006

Preliminary Draft: Please do not cite.

Abstract

Households in developing countries are frequently hit by severe idiosyncratic and covariate shocks resulting in high consumption volatility. A household's currently observed poverty status might therefore not be a good indicator of the household's general poverty risk, or in other words its vulnerability to poverty. Although several measurements to analyze vulnerability to poverty have recently been proposed, empirical studies are still rare as the data requirements for these measurements are not met by the surveys that are available for most developing countries. In this paper, we propose a simple method to empirically assess the impact of idiosyncratic and covariate shocks on households' vulnerability, which can be used in a wide context as it relies on commonly available living standard measurement surveys. We apply our approach to data from Madagascar and show, that whereas covariate shocks have a substantial impact on rural households' vulnerability, urban households' vulnerability is largely determined by idiosyncratic shocks.

JEL Classification: I32, D60.

Key words: Vulnerability to poverty, idiosyncratic and covariate shocks, multilevel modelling.

*University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany, email: isabel.guenther@wiwi.uni-goettingen.de.

†University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, email: k.harttgen@wiwi.uni-goettingen.de.

1 Introduction

Households in developing countries are frequently hit by severe idiosyncratic shocks (i.e. household-level shocks, such as death, injury or unemployment) and covariate shocks (i.e. community shocks, such as natural disasters or epidemics), resulting in high income volatility. Although households in risky environments have developed various sophisticated risk-coping strategies to reduce income fluctuations or to insure consumption against these income fluctuations, variance in household consumption remains generally high (see e.g. Townsend, 1994; Udry, 1995). A household's currently observed poverty status is therefore in many cases not a very good guide to the household's vulnerability to poverty, i.e. its general poverty risk. Whereas some households might be trapped into chronic poverty, others might only temporarily be poor, whereas other households currently non-poor might still face a high risk to fall into poverty in the future.

Most established poverty measurements, e.g. the FGT poverty measures (Foster et. al, 1984), do however only assess the current poverty status of a household without taking into account dynamic consumption fluctuations. Results from these static poverty analysis might therefore be misleading if high consumption volatility persists in a country. Not only might poverty rates fluctuate from one year to another, but even if aggregate poverty rates are constant over time, the share of the population which is vulnerable to poverty might be much higher. Moreover, these poverty measures cannot assess whether high poverty rates are a cause of structural poverty (i.e. low endowments) or a cause of poverty risk (i.e. high uninsured income fluctuations), which is important to know from a policy perspective.

To overcome these shortcoming of traditional poverty assessments, which can only present a static and ex-post picture of households' welfare, vulnerability analysis estimates the ex-ante welfare of households, taking into account the dynamic dimension of poverty. Vulnerability assessments therefore

try to estimate ex-ante both the expected mean as well as volatility of consumption, with the latter being determined by idiosyncratic and covariate shocks.

Although there has recently been a growing theoretical literature on vulnerability measurement, relevant empirical studies on vulnerability are - largely due to data limitations - still rare. First, to examine the dynamic aspects of poverty, lengthy panel data would be ideal. But for many developing countries, panel data does not exist and cross-sectional surveys are the only data available. Second, most household surveys were not designed to provide a full accounting of the impact of shocks. Information on idiosyncratic and covariate shocks is therefore in most data sets either completely missing or very limited. Hence, existing empirical studies have so far either only examined the aggregate vulnerability of households, ignoring the causes of the observed vulnerability, or have only studied the impact of selected idiosyncratic or covariate selected shocks on households' consumption, leaving out an analysis of the relative importance of different shocks on households' vulnerability as well as facing severe statistical problems.

The objective of this paper is hence to assess the relative impact of idiosyncratic and covariate shocks on households' vulnerability to poverty. More precisely, we both analyze how much of households' vulnerability is structural and risk induced, as well as provide an estimate of the share of consumption volatility that is idiosyncratic and covariate respectively. We propose a simple method which can be applied to commonly available standard household surveys without being constraint by the usual data limitations for vulnerability analysis; i.e. the impact of idiosyncratic and covariate shocks on households' vulnerability can be assessed without any panel and specific shock data. The proposed approach is an integration of multilevel analysis (Goldstein 1987, 1999) into the widely applied method by Chaudhuri (2000).

The remaining paper is structured as follows. Section 2 briefly discusses the current empirical literature on vulnerability to poverty, including its shortcomings. Section 3 proposes a methodology that allows assessing the relative importance of idiosyncratic and covariate shocks for households' vulnerability using cross-sectional data. Section 4 presents the data used and the estimation results. Section 5 concludes.

2 Empirics on Vulnerability to Poverty

As discussed in the introduction a household's currently observed poverty status might not be a reliable guide to a household's longer-term wellbeing. Policy makers and researchers in development economics have therefore long emphasized that it is critical to go beyond a static ex-post assessment of who is currently poor to a dynamic ex-ante assessment of who is vulnerable to poverty. But although there has been an emerging literature on both the theory and empirics of vulnerability, its significance especially for policy makers is still rather low.

The current state of the theoretical literature on vulnerability can be described in the words of Hoddinott and Quisumbing (2003) as a "let a hundred flowers bloom" phase of research with numerous definitions and measures and seemingly no consensus on how to estimate vulnerability. Several competing measurements have been offered (for an overview see e.g. Hoddinott and Quisumbing, 2003; Ligon and Schechter, 2004) and the literature has not yet settled on a preferred definition or measure. In this paper we use the measure proposed by Chaudhuri (2000) and Pritchett, Suryahadi and Sumarto (2000) who define vulnerability as expected poverty, or in other words as the probability that a household's consumption will lie below the poverty line in the near future.

But no matter how vulnerability is defined, i.e. which measurement of vulnerability is used, vulnerability is always a function of the expected mean

and variance of households' consumption, where the mean of expected consumption is determined by household and community characteristics whereas the variance in household consumption is determined by the occurrence and impact of idiosyncratic and covariate shocks as well as the strength of households' coping mechanisms to insure consumption against these shocks.

For a comprehensive understanding of vulnerability to poverty it would therefore be important to know both the magnitude of consumption volatility (i.e. the level of vulnerability) as well as what causes or reduces this volatility in consumption (i.e. the sources of vulnerability). Currently available data does however not even allow for a thorough assessment of the ex-ante vulnerability of households or the ex-post impact of shocks on consumption, let alone measure both the level and sources of vulnerability at the same time. The existing empirical literature is hence divided into two strands of literature; either concentrating on the measurement of aggregate vulnerability within a population or analyzing the ex-post impact of selected shocks on households' consumption.

The first strand of literature, which intends to estimate the aggregate vulnerability of households, has been pioneered by Townsend (1994) and Udry (1995), who were some of the first using panel data to analyze whether households are able to insure their consumption against idiosyncratic income fluctuations over space and time. In this spirit several studies followed analyzing consumption fluctuations over time (e.g. Dercon and Krishnan, 2000; Jalan and Ravallion, 1999; Morduch, 2005), concluding that households are partly but not fully capable of insuring consumption against income fluctuations. A severe drawback of this literature is that it relies on panel data, which is very limited for developing countries. The existing studies and drawn conclusions are hence often based on very few rounds (often not more than 3 waves) or observations (often not more than 100 households) of rural (urban households are mostly ignored) panel data (see also Morduch,

2005). A major confounding factor is here also the problem of measurement error as it is quite difficult to distinguish real consumption changes from measurement error in these relatively short panels (see e.g. Luttmer, 2001; Woolard and Klasen, 2005). However, in many developing countries panel data is completely missing and one has even to rely on cross-section surveys to estimate vulnerability.

The second strand of empirical literature on vulnerability, which estimates the impact of selected shocks on households' consumption, has also large (mostly) data-driven limitations. Information on idiosyncratic and covariate shocks is in most households surveys very limited and sometimes even completely missing (see also Günther and Harttgen, 2005). As a consequence most authors have only been able to focus on the impact of selected shocks on consumption (see e.g. Dercon and Krishnan, 2000; Gertler and Gruber, 2002; Glewwe and Hall, 1998; Kochar, 1995; Paxon, 1992). Concentrating on certain shocks does however not allow for an analysis of the relative impact of various shocks on households' consumption to assess which shocks should be given first priority in anti-poverty programs. Moreover, these studies have rarely been able to analyze the impact of these shocks on the vulnerability of households, as households' vulnerability to shocks is not only a function of the impact of shocks on households' consumption but also of the frequency distribution of these shocks.

In addition, there are severe econometric problems related to this work, which usually relies on standard regression analysis to study the impact of shocks on households' consumption. First, focusing on certain shocks introduces a considerable omitted variable bias as various shocks are often highly correlated (Mills et al, 2003; Tesliuc and Lindert, 2004). The impact of selected shocks on households' consumption is therefore likely to be overestimated. Second, it is often assumed that the impact of shocks on consumption is the same across all households, which is a rather strong

assumption to make. Third, the problem of endogeneity might be severe as households' welfare has presumably also an impact on the occurrence of certain shocks (e.g. poor households normally face higher mortality risks).

Last, several studies, which have analyzed the impact of covariate community shocks might have be biased by a disregard of the hierarchical data structure underlying these estimates.¹ If community shocks are simply assigned to each household within a community, blowing up data values from a small number of communities (upper-units) to many more household observations (sub-units), the assumption of independent observations is violated. However, ordinary statistical tests treat these disaggregated data values as independent information, leading to significant results that might be totally spurious (Steenbergen and Jones, 2002).

We certainly cannot bridge the data gaps that exist with regard to missing panel data and missing data on shocks in developing countries. What we propose is an approach, which allows to study the relative impact of idiosyncratic and covariate shocks on households' vulnerability, without any panel data and without facing the discussed econometric problems that usually occur when estimating the impact of certain shocks on household consumption. Furthermore, we estimate the level and sources of vulnerability simultaneously, which has rarely been done.

Although we cannot distinguish between the impact of individual shocks, a disaggregation into the impact of covariate community and idiosyncratic household specific shocks is already a valuable step forward. Idiosyncratic shocks are uncorrelated across households and should therefore be insurable by informal mutual insurance mechanisms within communities. Covariate shocks are correlated across households within the same community and informal insurance mechanism within communities should therefore break down during covariate shocks. And broad-base mutual insurance markets

¹We speak of hierarchical data structure whenever variables are measured at different hierarchical levels (see Goldstein, 1997, 1999), e.g. at the household and community level.

across communities do not function because of imperfect information and costly enforcement [...].

It is therefore claimed that only idiosyncratic risk can be insured within communities, where information and enforcement problems are less severe. Hence, analyzing the relative impact of covariate and idiosyncratic shocks on households' consumption first of all tests to what extent idiosyncratic shocks are really "less of a problem" than covariate shocks for households' consumption. Second, an assessment of the relative importance of idiosyncratic and covariate shocks might help policy makers to set up insurance priorities. Although higher information and enforcement problems prevail for insurance across communities, shocks that occur on the community level are easier to observe and also easier to mitigate with national safety nets as they are geographically clustered.

Few studies (see e.g. Carter, 1997; Dercon and Krishnan, 2000) have attempted to estimate the relative importance of covariate and idiosyncratic shocks on households' consumption. Their estimations generally show, that covariate shocks have a larger and more significant impact on households' consumption than idiosyncratic shocks . However, these studies have only analyzed rural households, relied on panel data, which is rarely available for developing countries and also faced the discussed econometric problems of concentrating on some selected idiosyncratic and covariate shocks, without taking into account the hierarchical data structure. In addition, it is often difficult to distinguish ex-ante between idiosyncratic and covariate shocks, as certain shocks often do have a covariate and idiosyncratic component [...]. Hence we think that our approach will contribute to a somewhat better understanding of the relative impact of idiosyncratic and covariate risks on households' vulnerability to poverty.

3 Methodology

3.1 The Mean and Variance of Consumption

Our proposed method is an extension of the methodology proposed by Chaudhuri (2002) to estimate expected mean and variance of consumption using cross-sectional data.² As for most developing countries panel data is not available this method which relies on only one cross-sectional survey has recently become quite popular [...]. The main hypothesis is that the error term in a cross-sectional consumption regression, or in other words the unexplained part of households' consumption, captures the impact of idiosyncratic and community specific covariate shocks, and that this cross-sectional variance also reflects inter-temporal variance in consumption. It is furthermore assumed that this variance in consumption can be explained by household and community characteristics, i.e. that the impact of shocks on consumption fluctuations is correlated with observable variables.

Suppose that a household's h consumption in period t is determined by a set of variables X_h . We can hence set up the equation

$$\ln c_h = X_h \beta + e_h \tag{1}$$

where $\ln c_h$ is the log of per capita household consumption, X_h a set of household as well as community characteristics, and e_h the part of households' consumption that cannot be explained. Chaudhuri, Jalan and Suryahadi (2002) suggest that this error term, or the variance in consumption of otherwise equal households, captures the impact of both idiosyncratic and community specific covariate shocks on households' consumption and that this variance is correlated with observable household and community characteristics. In a second step, the variance of the error term is therefore regressed on the same (and other) household and community characteristics

²For a detailed discussion of the method see Chaudhuri (2002) and Chaudhuri, Jalan and Suryahadi (2002).

$$\sigma_{eh}^2 = X_h \theta. \quad (2)$$

Standard regression analysis using ordinary least squares (OLS) estimation techniques assume homoscedasticity, i.e. the same variance $V(e_i) = \sigma^2$ across all households i . However, as discussed, Chaudhuri (2000) assumes that the variance of the error term is not equal across households, reflecting the impact of shocks on consumption; i.e. the error term is assumed to be heteroscedastic. Using OLS for an estimation of β and θ would therefore lead to unbiased but inefficient coefficients.

To overcome the problem of heteroscedasticity, equation (1) has to be reduced to a model where the residuals e_h have a homogeneous variance (for a detailed discussion see Maddala, 1977). Chaudhuri (2000) hence applies three-step feasible generalized least squares (FGLS) to estimate efficient coefficients β and θ . In principle this means, first estimating (1) via OLS, then estimating (2) again by OLS using the squared residuals of (1) as the dependent variable. The predictions from (2) are then used to weight and reestimate (2). In a last step the now efficient coefficients θ can be used to predict again (2), which is then used to weight equation (1) and reestimate it, obtaining also efficient estimates for β (see Chaudhuri, 2000 or Chaudhuri, Jalan and Subyahadi, 2002 for a detailed discussion of the methodology).

In a third step, for each household the expected mean as well as variance of consumption can be estimated using the consistent and asymptotically efficient estimators $\hat{\beta}_{FGLS}$ and $\hat{\theta}_{FGLS}$.

$$\hat{E}[lnc_h | X_h] = X_h \hat{\beta} \quad (3)$$

$$\hat{V}[lnc_h | X_h] = \hat{\sigma}_{e,h}^2 = X_h \hat{\theta}. \quad (4)$$

In the absence of any information on time-variant consumption volatility in a cross-sectional survey, two rather strong assumptions have to be made

when applying this approach. First, it is assumed that cross-sectional variance can be used to estimate inter-temporal variance in consumption. Certainly, cross-sectional variance can explain part of inter-temporal variance due to idiosyncratic or covariate community-specific shocks. However, the model will miss the impact of inter-temporal shocks on the national level (for example terms of trade shocks). Second, it is hypothesized that the impact of shocks on consumption variance is correlated with household characteristics, whereas measurement error is not correlated with household characteristics.

However, the proposed method has the great advantage that it overcomes both the problem of missing panel as well as incomplete information on shocks, which might often lead to biased results with regard to the impact of shocks on households' consumption.

We extend the proposed method by Chaudhuri (2000), introducing multilevel analysis (Goldstein, 1999). This first of all allows us to differentiate between the unexplained variance on the household level (i.e. the impact of idiosyncratic shocks) and the unexplained variance on the community level (i.e. the impact of covariate shocks). Second, multilevel analysis corrects for inefficient estimators, which might occur whenever the proposed methodology by Chaudhuri (2002) is applied to hierarchical data structures, i.e. whenever variables from various levels (e.g. from the household and community level) are introduced in the regressions.

3.2 Multilevel Analysis

Multilevel models are designed to analyze the relationship between variables that are measured at different hierarchical levels (see e.g., Bryk and Raudenbush, 1992; Goldstein, 1987, 1999; and Hox, 2002). Thus, multilevel analysis explicitly takes into account hierarchical data structure and allows to use both household and community variables simultaneously in the same model without violating the assumption of independent observations, pro-

viding correct standard errors and significance tests (Goldstein 1999). In addition, multilevel models allow for a decomposition of the error term; this means for our case decomposing the unexplained variance of consumption into a household and community component.

To illustrate the basic idea of multilevel modelling suppose $i = 1, \dots, n_i$ level one units (households) and $j = 1, \dots, n_j$ level two units (communities) and that the household i is nested within the community j . If Y_{ij} is (in our case) per capita household consumption and X_{ij} a set of household characteristics of household i in community j then we can set up a regression equation as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (5)$$

where the error term e_{ij} reflects the unexplained part in households' consumption. Note that in contrast to standard regression models, the variables in equation (5) are denoted by two subscripts: one referring to the household i and one to the community j , and that the coefficients are denoted by a subscript referring to the community j . This means that it is assumed that β_{0j} and β_{1j} vary across communities. Various community characteristics Z can then be introduced into the model to estimate the variance of the coefficients across communities.

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (6)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}. \quad (7)$$

where the error terms u_{0j} and u_{1j} represent level two residuals, i.e. the unexplained variance in consumption between communities.³ Equation (7) and (8) hence reflect the impact of community characteristics Z on household

³The residuals u_{0j} and u_{1j} are assumed to have a mean of zero, $E(u_{0j}) = E(u_{1j}) = 0$. The variance of u_{0j} and u_{1j} is $var(u_{0j}) = \sigma_{u0}^2$ and $var(u_{1j}) = \sigma_{u1}^2$ respectively, and the covariance is $cov(u_{0j}, u_{1j}) = \sigma_{u01}$. It is furthermore assumed that level one residuals are not correlated with level two residuals, $cov(u_{0j}, e_{ij}) = cov(u_{1j}, e_{ij}) = 0$.

consumption which differs across communities but which is the same for households within the same community j .

Substituting equation (6) and (7) into equation (5) provides the full model, which can be written as:

$$Y_{ij} = \overbrace{\gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + \gamma_{11}X_{ij}Z_j}^{\text{deterministic}} + \overbrace{(u_{0j} + u_{1j}X_{ij} + e_{ij})}^{\text{stochastic}}. \quad (8)$$

and estimated via maximum likelihood (Mason et al, 1983; Goldstein, 1987; Bryk and Raudenbush, 1992).⁴ The first part of equation (8) reflects the deterministic part of the equation, including the interaction term $X_{ij}Z_j$, which analyzes cross-level interactions between variables at the household and variables at the community level. The second part, expressed in brackets, captures the stochastic part of the model. In contrast to standard OLS regression the error term in (8) contains not only an individual or household component e_{ij} but also a group or community component $u_{0j} + u_{1j}X_{ij}$. The error term u_{0j} represents the unexplained variance across communities for the intercept β_{0j} . The error term u_{1j} reflects the unexplained variance across communities for the slopes β_{1j} . The error term e_{ij} captures the remaining unexplained individual or household variance in consumption.

The stochastic part in equation (8) demonstrates the problem of dependent errors in multilevel data structure. Whereas the household error component e_{ij} is independent across all households, the community level errors u_{0j} and u_{1j} are independent between communities but dependent within each community, as the error terms are equal for every household i within community j . This also leads to heteroscedastic error terms, as the error term of a household depends on u_{0j} and u_{1j} which vary across communities and on household characteristics X_{ij} which vary across households.

⁴In a more general form, assuming P explanatory variables X at the lowest level, denoted by the subscript $p(p = 1...P)$ and Q explanatory variables Z at the highest level, denoted by the subscript $q(q = 1...Q)$ the equation is $Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + \gamma_{pq}X_{pij}Z_{qj} + (u_{pj}X_{pij} + u_{0j} + e_{ij})$.

3.3 The Impact of Idiosyncratic and Covariate Shocks

To assess households' vulnerability to idiosyncratic and covariate shocks using cross-sectional data we extend the method of Chaudhuri (2000) by incorporating multilevel modelling. In a first step we regress the log of per capita household consumption of household i in community j on a set of household X and community covariates Z using a basic two level model.

$$\ln c_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + (u_{0j} + e_{ij}). \quad (9)$$

The difference to equation (8) is that in equation (9) no cross-level interactions are included so that the interaction term $X_{ij}Z_j$ and the error part $u_{1j}X_{ij}$ are set to zero⁵. Equation (9) hence estimates two error terms u_{0j} and e_{ij} . Following Chaudhuri (2000) it is supposed that the error term at the household level e_{ij} captures the impact of idiosyncratic shocks whereas the error term at the community level u_{0j} captures the impact of covariate shocks on households' consumption.

In a second step we then estimate the variance at the household level ($\sigma_{e_{ij}}^2$) and the community level ($\sigma_{u_j}^2$) using the squared residuals from equation (9), again applying multilevel analysis, which provides us with asymptotically efficient and consistent estimation parameters for each variance component.

$$\sigma_{e_{ij}}^2 = X_{ij}\theta_1 + Z_j\theta_2 \quad (10)$$

$$\sigma_{u_j}^2 = Z_j\theta_3 \quad (11)$$

In a third step we predict the mean (12) as well as the variance of households' consumption that is caused by idiosyncratic (13) and covariate shocks

⁵The usual procedure for multilevel modelling is to build up the model in several steps. The outset is a model with only level one (household) variables as a benchmark model. Then higher level (communities) variables are included (Hox, 2002), but without any cross-level interaction effects. In a last step interaction terms are included. Incorporating interaction terms and the set-up of a full multilevel model is left for a later version of the paper

(14):

$$lnc_{ij} = X_{ij}\hat{\gamma}_{10} + Z_j\hat{\gamma}_{01}. \quad (12)$$

$$\hat{V}_{idiosyncratic}[lnc_{ij}|X_{ij}, Z_j] = \hat{\sigma}_{eij}^2 = X_{ij}\hat{\theta}_1 + Z_j\hat{\theta}_2 \quad (13)$$

$$\hat{V}_{covariate}[lnc_{ij}|Z_j] = \hat{\sigma}_{uj}^2 = Z_j\hat{\theta}_3. \quad (14)$$

Based on the estimated mean and variance of consumption any measure of vulnerability can be applied to assess the impact of idiosyncratic and covariate shocks on households' vulnerability.

4 Empirical Illustration

4.1 Data and Model Specification

We empirically illustrate our proposed approach for Madagascar. Madagascar is one of the poorest countries in Sub-Saharan Africa with a GDP per capita of 744 USD PPP and an estimated headcount poverty rate of about 70 percent (World Bank, 2005). Its poor economic performance is accompanied by very low social indicators of human well-being. Life expectancy at birth is 55 years and high rates of child mortality [...] and child undernutrition [...] persist.

Moreover, households in Madagascar are frequently hit by idiosyncratic and covariate shocks which have an additional severe down-side impact on households' well-being (Mills, Ninno and Rjemison 2003). Mills, Ninno and Rjemison (2003) report that households are most notably hit by frequently occurring covariate shocks, in particular climatic shocks like droughts and cyclones, which also show a quite strong spatial and temporal correlation (Mills, Ninno and Rjemison 2003).

The data which we use for our analysis is derived from a cross-sectional household survey and a cross-sectional community census. The community

census is the 2001 ILO/Cornell Commune Levels census which provides information on community characteristics like social and economic infrastructure as well as data on the occurrence of covariate shocks. It covers 1,385 out of the 1,395 communities in Madagascar. Data on household characteristics is taken from the national representative household survey of 2001 (Enquete Aupres Des Menages, EPM), covering 5,080 households in 180 communities.

To estimate households' expected mean and variance of consumption we include a set of household and community characteristics in our model (see Table 1). In addition to the household characteristics listed in Table 1, we consider an household asset index estimated via principal component analysis (see Filmer and Pritchett, 2001) on several agricultural assets. At the community level we include several variables reflecting the social and market infrastructure of the communities as well as their population density, which might also influence households' consumption. The community characteristics do not enter separately into the model but as an infrastructure index based again on a principal component analysis.

Table 1
Summary statistics for household and community characteristics
for Madagascar (2001)

	Urban	Rural	National
Household characteristics			
Age of household head (in years)	42.60	41.71	42.25
Sex of household head (1=male)	76.70	78.07	77.60
Education of household head (in years)	7.80	4.15	6.35
Household size	4.42	4.78	4.56
Total no. children	1.70	2.16	1.88
Number of cattle	0.93	4.88	2.50
Number of chicken	2.63	8.70	5.04
Working in informal sector (%)	22.88	7.04	16.59
Working in formal sector (%)	21.74	5.80	15.41
Working in agricultural sector (%)	41.02	83.00	57.68
Employed (%)	43.86	57.27	49.19
Households having an enterprize in the non-agricultural sector (%)	30.22	20.24	26.26
Community characteristics			
Telephone (%)	83.16	18.75	57.60
Sanitation (%)	75.26	20.54	53.54
Save water (%)	98.43	50.00	79.21
Electricity (%)	98.43	42.00	76.02
Primary education (%)	100	100	100
Secondary education (%)	100	67.86	87.16
Tertiary education (%)	97.89	10.71	63.07
Hospital (%)	93.01	7.14	58.53
National road (%)	93.67	53.75	77.65

Source: Own calculations using the 2001 Enquete Aupres Des Menages (EPM) and 2001 ILO/Cornell Commune Levels census.

4.2 Estimation Results

As described in section 3 we first estimate the expected mean and variance of log per capita consumption using multi level modelling. We furthermore decompose unexplained variance in consumption into an idiosyncratic (household-level) and covariate (community-level) component. To remind, we assume that the estimated variance in consumption on the household-level reflects the impact of idiosyncratic shocks on household consumption whereas the estimated variance in consumption on the community-level reflects the impact of covariate shocks on household consumption. In many

studies the village has been used as the "natural" covariate shock (and mutual insurance) level, but there is no necessity to do so (Genicot and Ray, 2003; Morduch, 2005), and using communities instead, as we do in this analysis, seems not much less useful.

Estimation results are presented in Table 2 separately for rural and urban households, representing 69 percent and 31 percent of national households respectively. The expected per capita (log) consumption of rural households is considerably below the (log) the poverty line, whereas the expected per capita (log) consumption of urban households lies considerably above this line. This already indicates that low mean consumption is the main cause for rural vulnerability.

Table 2
Estimated mean and variance of consumption
for Madagascar (2001)

	Rural	Urban	National
Households	0.69	0.31	1.00
Consumption, <i>Inc</i> (predicted)			
(ln) per capita expenditure	13.54	14.25	13.76
(ln) poverty line	13.80	13.80	13.80
Standard deviation (predicted)			
Standard deviation (total)	0.60	0.60	0.60
Standard deviation (idiosyncratic)	0.47	0.53	0.49
Standard deviation (covariate)	0.33	0.25	0.31

Source: Own calculations using the 2001 Enquete Aupres Des Menages (EPM) and 2001 ILO/Cornell Commune Levels census.

Note: Values are household weighted.

With regard to the estimated mean variance in consumption, we show that whereas total estimated variance is the same for rural and urban households, with a standard deviation of 0.60 (see Table 2), the relative importance of idiosyncratic variance is much higher for urban than for rural households. More precisely, whereas among urban households the idiosyncratic standard deviation of consumption is 2.21 as high as the covariate standard deviation, the respective rate is only 1.41 for rural households. This denotes that

idiosyncratic shocks have a relatively high impact on urban consumption whereas covariate shocks have a relatively higher impact on rural consumption. Interesting to note is also that the average idiosyncratic variance seems to be higher than covariate variance in households' consumption for both rural and urban households. In addition to Table 2, which presents the mean of variance in consumption, Figure 1 also shows the distribution of the covariate and idiosyncratic variance in consumption across urban and rural households. [...]

[Figure 1]

To obtain a full assessment of the level and sources of vulnerability, we have to assess expected mean and variance of consumption jointly across the consumption distribution. Although various vulnerability measurements could be applied here, for this study we only use the definition proposed by Chaudhuri, Jalan and Suryahadi (2002), assessing vulnerability to poverty as the probability of a household to fall below the poverty line in the near future. Assuming that consumption is log-normally distributed, we can estimate the probability of a household to fall below the poverty line using the estimated expected mean (equation (12)) and variance (equation (13) and (15)) of consumption.

$$\hat{v}_h = \hat{P}(\ln c_h < \ln z | X_h) = \Phi \left(\frac{\ln z - X_h \hat{\beta}}{\sqrt{X_h \hat{\theta}}} \right) \quad (15)$$

where $\Phi(\cdot)$ denotes the cumulative density of the standard normal distribution function and z denotes the poverty line.

Last, we have to define a probability or vulnerability threshold above which we consider households as vulnerable to poverty as well as the time horizon which we consider as the "near future". In this study vulnerability to poverty is defined as a 50 percent or higher probability to fall below the

poverty line.⁶ The time horizon we apply is $t+2$ years. This means, that we consider those households as vulnerable which have a 50 percent or higher probability to fall below the poverty line at least once in the next two years, which is equivalent to a 29 percent or higher probability to fall below the poverty line. Or in other words, households that have a 50 percent or higher probability to fall below the poverty line at least once in the next two years, must have a 29 percent probability or higher to fall below the poverty line in the next period.⁷

Table 3
Vulnerability decomposition
in Madagascar (2001)

	Rural	Urban	National
Total Vulnerability	0.91	0.48	0.76
Low mean	0.68	0.11	0.49
High volatility	0.23	0.37	0.27
Idiosyncratic Vulnerability	0.87	0.44	0.72
Low mean	0.68	0.11	0.49
High volatility	0.19	0.33	0.23
Covariate Vulnerability	0.84	0.33	0.66
Low mean	0.68	0.11	0.49
High volatility	0.16	0.22	0.17

Source: Own calculations using the 2001 Enquete Aupres Des Menages (EPM) and 2001 ILO/Cornell Commune Levels census.

Note: Values are household weighted.

Utilizing the stated vulnerability threshold and time horizon we estimate that 76 percent of households in Madagascar are vulnerable to poverty within the next two years (Table 3). The respective figures for urban and rural households are 91 and 48 percent respectively, indicating that (as expected) rural households are much more vulnerable to poverty than urban households.

⁶The 50 percent threshold has become a standard vulnerability threshold in the literature [see e.g. ...].

⁷For a detailed discussion see Günther and Harttgen, 2005.

We furthermore decompose these vulnerability estimates into sources of vulnerability. In other words we first analyze whether vulnerability is mainly driven by permanent low consumption prospects (i.e. structural poverty) or by high consumption volatility (i.e. high poverty risk). We state that rural vulnerability is mainly a cause of low expected mean in consumption whereas urban vulnerability is mainly driven by high consumption volatility (Table 3). More precisely, 68 percent of rural households have an expected per capita consumption that already lies below the poverty line, and "only" 23 percent of the 91 percent vulnerable rural households are vulnerable because of high consumption volatility. In contrast, 37 percent of urban households face risk induced vulnerability (i.e. high consumption fluctuations) whereas only 11 percent face structural induced vulnerability.

Last, we analyze the impact of idiosyncratic and covariate shocks on vulnerability to poverty. As already indicated in Table 2 and Figure 1 idiosyncratic shocks have a slightly higher influence than covariate shocks on consumption among rural households and a much higher influence than covariate shocks on households' consumption in urban areas (Table 3). 87 percent of rural and 44 percent of urban households are vulnerable to idiosyncratic shocks whereas 84 percent of rural and "only" 33 percent of urban households are vulnerable to covariate shocks.

As an assessment of vulnerability to poverty depends not only on the poverty line but also highly on the chosen vulnerability or probability threshold above which we consider households as being vulnerable to poverty, we also show the cumulative density distribution of vulnerability to poverty in Figure 2. It presents the percentage of households that have a i or higher probability to fall below the poverty line. Again, estimates are provided for Madagascar as a whole and for rural and urban households separately.

In Figure 2, we marked the vulnerability threshold of 29 percent, which we used for our vulnerability analysis, providing us with the same estimates

as presented in Table 3. What is now interesting to see is, that the relative importance of covariate and idiosyncratic shocks for rural and urban households' consumption depends on the vulnerability threshold chosen. Moreover, if we regard the whole cumulative density distribution of vulnerability to poverty, we observe that the share of urban households that face an idiosyncratic shock induced vulnerability is larger than the share of households that face a covariate shock induced vulnerability for the major part of vulnerability thresholds (Figure 2(b)), whereas the contrary is true for rural households, where covariate shocks seem to be more important for most vulnerability thresholds (Figure 2(a)). [...]

[Figure 2]

5 Conclusion

We proposed a simple method to analyze the level and sources of vulnerability using currently available standard cross-sectional households surveys without any explicit information about idiosyncratic and covariate shocks. In particular, the suggested method allows to estimate expected mean and variance in consumption of households, decomposing variance in consumption into an idiosyncratic and covariate part.

Using the concept of Chaudhuri (2000), defining vulnerability to poverty as the probability of a household to fall below the poverty line, we stated that both covariate and idiosyncratic shocks have a considerable impact on both urban and rural vulnerability. Furthermore, our results indicate that idiosyncratic shocks seem to have an even higher impact on households' consumption volatility than covariate shocks. However, idiosyncratic shocks seem to have a relatively higher impact on urban households' consumption and covariate shocks seem to have a relatively higher impact on rural households' vulnerability.

It is difficult to say whether a higher impact of certain types of shocks

on rural or urban households' consumption is the result of a more severe impact of these shocks on households' income or the result of worse insurance mechanisms of certain households against these shocks. In the following we still provide some cautious explanations for our estimates.

The suggested high impact of idiosyncratic shocks on both rural and urban consumption implies that either insurance mechanisms within communities do not function any better than insurance mechanisms across communities or that idiosyncratic shocks have a much higher impact on households' income than covariate shocks, for example because covariate shocks might in many cases be more anticipated than idiosyncratic shocks, so that ex-ante coping strategies can be implemented.

The relatively higher impact of covariate shocks on rural households consumption might be explained by the fact that there are certainly many more covariate shocks (such as climatic shocks) which have a higher impact on rural (agricultural) households than on urban (non-agricultural) households. Also, it is possible that urban households face even higher information and enforcement problems and that therefore mutual community based informal insurance mechanisms work better among rural than among urban households, mitigating the adverse effects of idiosyncratic shocks in rural but not in urban areas. Rural households might also have better self-insurance mechanisms in place. [...]

Last, we noted that the relative importance of consumption fluctuations (versus low mean consumption) seems to be even greater for urban households' welfare than for rural households' welfare. Hence, urban households should - if possible - be included into vulnerability studies, which have so far mostly focused on rural villages and households.

We are aware of the fact, that some rather stringent assumptions have to be made to apply the proposed method. However, we argue that as long as lengthy panel data with comprehensive information on idiosyncratic

and covariate shocks is missing, the suggested approach can provide quite interesting insights into the impact of idiosyncratic and covariate shocks on households' vulnerability. Moreover, we recommend, that any study which analyzes the influence of covariate shocks on households' consumption - no matter if cross-sectional or panel-data is used and independent of the extent of shock data available - should apply multilevel modelling as it appropriately takes into account the hierarchical structure of the data that is used for such analysis.

References

- Bryk, A.S. and S.W. Raudenbush (1992), *Hierarchical Linear Models: Applications and Data Analysis*, Newbury Park, CA: Sage Publications.
- Carter M. (1997), Environment, Technology, and the Social Articulation of Risk in West African Agriculture *Economic Development and Cultural Change*, 45 (3): 557-590.
- Chaudhuri, S. (2000), Empirical methods for assessing household vulnerability to poverty, mimeo, Department of Economics, Columbia University, New York.
- Chaudhuri, S., J. Jalan and A. Suryahadi (2002), Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia, Department of Economics Discussion Paper No. 0102-52, Columbia University, New York.
- Dercon S. and P. Krishnan (2000), Vulnerability, seasonality and poverty in Ethiopia. *Journal of Development Studies*, 36 (6): 25-53.
- Filmer, D. and L.H. Pritchett (2001), Estimating Wealth Effects without Expenditure Data - or Tears: An Application to Educational Enrollments in States of India, *Demography* 38 (1): 115-132.
- Foster, J., Greer, J. and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica* 52: 761-5.
- Genicot, G. and D. Ray (2003), Endogenous Group Formation in Risk-Sharing Arrangements, *Review of Economic Studies*, 70(1): 87-113.
- Gertler P. and J. Gruber (2002), Insuring Consumption Against Illness. *American Economic Review* 92(1): 55-70.

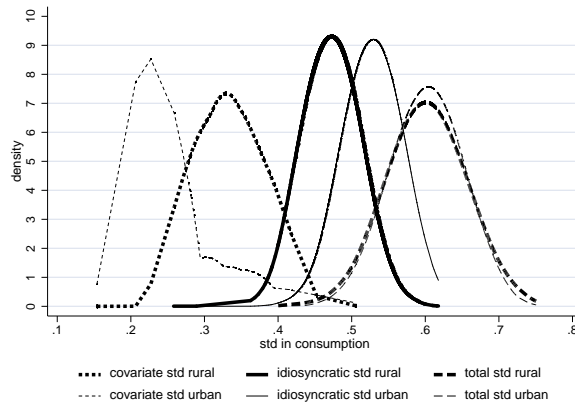
- Glewwe P. and G. Hall. (1998). Are some groups more vulnerable to macroeconomic shocks than others. Hypothesis tests based on panel data from Peru. *Journal of Development Economics* 56 (1): 181-206.
- Goldstein, H. (1987), *Multilevel Models in Educational and Social Research*, London, Griffin.
- Goldstein, H. (1999), *Multilevel Statistical Models*, London, Arnold.
- Gorad, S. (2003), What is Multi-Level for?, *British Journal of Educational Studies*, 51 (1): 46-63.
- Günther I. and K. Harttgen (2005), Vulnerability to Poverty in Madagascar. Background Paper prepared for the World Bank Social Protection Unit.
- Hoddinott J. and A. Quisumbing (2003), Methods for Microeconometric Risk and Vulnerability Assessments. *Social Protection Discussion Paper* 0324, World Bank, Washington.
- Hox, J. (2002), *Multilevel Analysis - Techniques and Applications*, Mahwah, NJ: Erlbaum.
- Jalan J. and M. Ravallion (1999), Are the poor less well insured? Evidence on vulnerability to income risk in rural China. *Journal of Development Economics* 58(1): 61-81.
- Kochar A. (1995), Explaining Household Vulnerability to Idiosyncratic Income Shocks. *The American Economic Review* 85(2): 159-164.
- Ligon E. and S. Schechter (2004), Evaluating Different Approaches to Estimating Vulnerability, *Social Protection Discussion Paper Series*, No. 0410, World Bank, Washington.

- Luttmer E.F.P. (2000), Inequality and Poverty Dynamics in Transition Economies: Disentangling Real Events from Noisy Data. Mimeo, World Bank, Washington.
- Madalla, G.S. (1977), *Econometrics*, New York.
- Mason, W.M., Wong, G.M. and B. Entwistle (1983), Contextual analysis through the multilevel linear model, in S. Leinhardt (ed.) *Sociological Methodology*, San Francisco, Jossey-Bass.
- Mills B., Ninno, C. and H. Rajemison (2003), Commune Shocks, Household Assets and Economic Well-Being in Madagascar, mimeo, World Bank, Washington.
- Morduch J. (2005), Consumption smoothing across space: Testing theories of risk-sharing in the ICRISAT study region of South India, In: *Insurance against Poverty*, ed. S. Dercon, Oxford University Press, Oxford.
- Paxon C.H. (1992), Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand. *The American Economic Review* 82(1): 15-33.
- Pritchett L., Suryahadi A. and S. Sumarto (2000), Quantifying Vulnerability to Poverty: A Proposed Measure, with Application to Indonesia. *SMERU Working Paper*, Social Monitoring and Early Response Unit (SMERU), World Bank, Washington.
- Sahn, D.E. and D. Stiefel (2003), Exploring Alternative Measures of Welfare in the Absence of Expenditure Data, *Review of Income and Wealth* 49 (4): 463-489.
- Steenbergen, M.R. and B.S. Jones (2002), Modelling Multilevel Data Structures, *American Journal of Political Science*, Vol. 46 (1): 218-237.

- Tesliuc E. and K. Lindert (2004), Risk and Vulnerability in Guatemala: A Quantitative and Qualitative Assessment. *Social Protection Discussion Paper* 0404, World Bank, Washington.
- Townsend R.M. (1994), Risk and Insurance in Village India, *Econometrica* 62 (3): 539-591.
- Udry C. (1995), Risk and Saving in Northern Nigeria, *The American Economic Review* 85 (5): 1287-1300.
- Woolard I. and S. Klasen (2005), Income Mobility and Household Poverty Dynamics in South Africa. *Journal of Development Studies* (forthcoming).
- World Bank (2005), *World Development Report 2005*, World Bank, Washington.

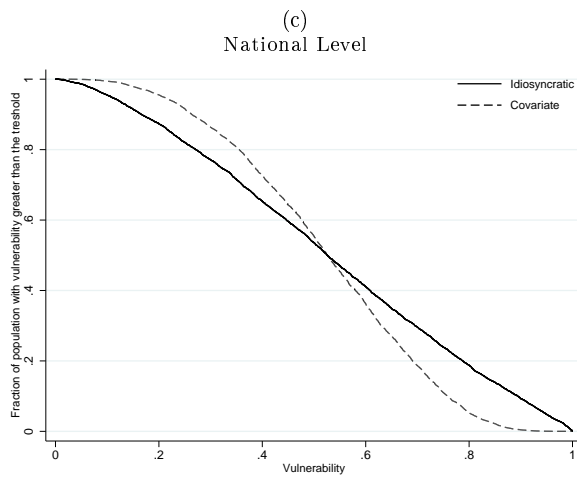
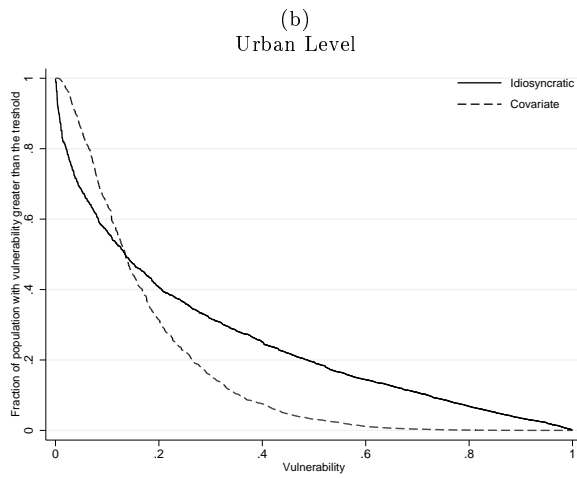
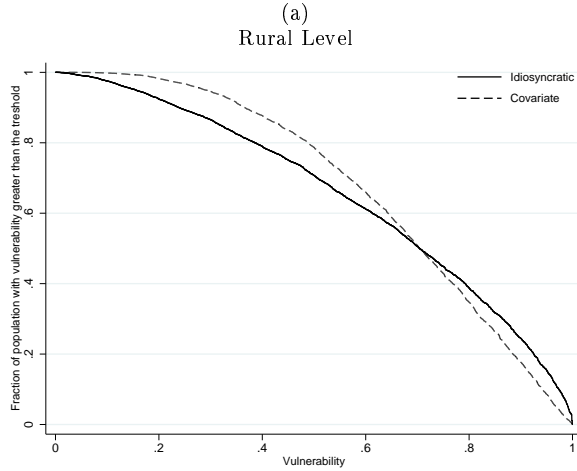
Figures

Figure 1
Density Distribution of Estimated Standard Deviation of Consumption



Source: Own calculations.

Figure 2
Cumulative Density Distribution of Vulnerability to Poverty



Source: Own calculations.