Vulnerability in a Stochastic Dynamic Model

Chris Elbers

Jan Willem Gunning*

Department of Economics, Vrije Universiteit, Amsterdam.

* Tinbergen Institute.

Vulnerability in a Stochastic Dynamic Model

Chris Elbers and Jan Willem Gunning[1]

Free University, Amsterdam

Revised September 2003

Key words: vulnerability, expected poverty, risk, Ramsey model, consumption regressions

JEL codes: D12, D60, D91, O12

Abstract

Most measures of vulnerability are a-theoretic and essentially static. In this paper we use a stochastic Ramsey model to find a household's optimal welfare and we measure vulnerability as the shortfall from the welfare attained if the household consumed permanently at the poverty line. The results indicate that vulnerability is very sensitive to the time horizon considered. We find that the accuracy of existing regression-based vulnerability measures can be greatly improved by including asset measures in the regression.

1   Introduction

It has long been recognised that a substantial part of poverty in developing countries is transient rather than chronic. Baulch and Hoddinott (2000) survey thirteen panel data studies of "poverty dynamics" - movements in and out of poverty - and conclude (p. 6) that "[i]n most of the studies, the category of 'sometimes poor' is larger, sometimes by a considerable amount, than the 'always poor'." Clearly, in such circumstances it may be very misleading to identify a household's normal poverty status on the basis of one-off survey data. A household with a permanent income well above the poverty line might appear to be poor if it was observed just after experiencing an unfavourable shock. Conversely, a household which in most years experiences poverty could be misclassified as non-poor if its income was observed just after a positive shock. Establishing the extent of transient poverty is important since chronic and transient poverty have, obviously, very different policy implications. The problem is only slightly less serious if poverty measures are based on consumption rather than income: typically capital market imperfections severely constrain a household's scope for consumption smoothing (*e.g.* Deaton, 1990).

That a household's current poverty[1] may be a bad guide to its future welfare explains the recent emphasis in the poverty literature on vulnerability (*e.g.* World Bank, 2001), a forward-looking concept. Vulnerability is often understood as the *expected* poverty of a household. Vulnerability is then calculated as poverty at some future date for all possible realizations of income or consumption, weighted by the probability of these outcomes. For the headcount poverty measure this implies that vulnerability is measured as the probability that the household will find itself below the poverty line at the specified date.

---

[1] In this paper we define a household's poverty status on the basis of the household's income or consumption. Hence we are concerned with income (or consumption) poverty.

Vulnerability is the net effect of three processes. It reflects, *first*, non-stochastic poverty determinants such as the soil quality of the holding or the education of the household's members, *secondly*, the household's exposure to shocks (*e.g.* unreliable rainfall) and, *thirdly*, its ability to cope with shocks (*e.g.* through insurance or the use of savings for consumption smoothing). The literature does distinguish between the non-stochastic and stochastic determinants of vulnerability but it does not make the further distinction between a household's exposure to risk and is ability to cope with risk.

To apply the concept of vulnerability empirically one must estimate the distribution of the household's consumption at some future date. One approach in the literature is to assume that all households face the same, stationary distribution so that the distribution can be estimated from cross-section data. A second method is to allow for inter-household heterogeneity. Maintaining the stationarity assumption one can then estimate household-specific distributions from time series data. These methods have the virtue of simplicity but, obviously, can lead to very misleading results if the stationarity assumption or, in the first case, the assumption of homogeneity are not satisfied. A third method is to regress a household's consumption on household characteristics and measures of realized shocks (*e.g.* an illness in the household). The estimated coefficients can then be used to predict the household's poverty for a particular (essentially arbitrary) realization of shocks. Households with high predicted poverty are then considered as vulnerable conditional on these shocks.[2] None of the existing measures is based on a structural model although some authors are well aware that vulnerability measures should be based on a dynamic analysis and that this "would require the structure of a proper dynamic

---

[2]This method, used by Dercon and Krishnan (2000), does not take into account the probability of the realization considered and thereby fails to capture the extent to which the household is exposed to shocks. Clearly, the result cannot be interpreted as a measure of *expected* poverty since no information on the distribution of future consumption is used.

model" (Ligon and Schechter, 2003, p. C101).

In this paper we argue that vulnerability measures currently in use are unsatisfactory: the expected poverty concept has unappealing characteristics (*e.g.* when household welfare increases - under a plausible definition of welfare - then expected poverty may rise) and the regression-based methods are likely to miss a large part of the impact of risk on household welfare. We propose an alternative methodology. The key step is to estimate a structural dynamic model of the household's consumption and (dis)saving, modelled as the outcome of intertemporal optimisation under uncertainty. This ensures that the household's responses (both *ex ante* and *ex post*) to shocks are explicitly taken into account. The estimated model can then be used to derive simulation-based estimates of vulnerability on the basis of a welfare concept, *viz.* expected discounted utility. We illustrate this method with an example, using the parameter estimates of Elbers *et al.* (2003) who estimate a stochastic Ramsey model on panel data for smallholder households in Zimbabwe.

We use this example to illustrate that vulnerability can change dramatically over time (both as a result of sustained growth and as a result of adjustment to shocks) so that outcomes are quite sensitive to the choice of time horizon. These results suggest that without a structural model vulnerability measures can be seriously misleading. We also show that much of the effect of risk on the mean of the ergodic distribution of consumption reflects the *ex ante* effect. The implication is that the usual identification of chronic poverty with structural determinants and transient poverty with risk breaks down: a household can be chronically poor *because* its response to risk lowers consumption permanently. Existing vulnerability measures fail to take this into account. Our results suggest that they thereby vastly underestimate the value of policies which reduce risk or improve households' ability to cope with risk. Such policies do not only reduce the volatility of household consumption around a given mean, but they also (and much more importantly) lower vulnerability by raising that mean.

3

In many situations the data required for a stochastic Ramsey model will not be available. We use our simulation estimates to assess the accuracy of less data-intensive regression-based vulnerability measures. Our key finding is that accuracy can be greatly improved if asset ownership (in our case: livestock) is included in the regression. Regressions which relate consumption to household characteristics (such as education or household size) and (*ex post*) shock measures but not to assets can be seriously misleading in identifying vulnerable households.

The structure of the paper is as follows. In the next section we review the methodology of vulnerability measures and we propose a definition which is explicitly dynamic. In section 3 we present the stochastic Ramsey model estimated by Elbers *et al.* (2003). In section 4 we compare the vulnerability estimates generated by this model and the measures derived from commonly-used regression specifications. Section 5 concludes.

2 Vulnerability Measures

When vulnerability is defined as expected poverty at a specified future date t (*e.g.* Christiaensen and Subbarao, 2001) it may be measured as

$$V = \int_0^z p(c, z) dF(c) \tag{1}$$

where z is a poverty line, c consumption, $F(c)$ the distribution of consumption at t and $p(c, z)$ a poverty measure, *e.g.* a member of the Foster-Greer-Thorbecke class

$$p(c, z) = \left[\frac{\max(z - c, 0)}{z} \right]^\alpha$$

where α is a non-negative parameter. The distribution F is taken as given and reflects both the household's exposure to shocks (idiosyncratic or covariant) and its ability to cope with them. In that sense F is a reduced form. (In the next section we will relax the assumption that F is

4

given and assume instead that only the distribution of shocks is exogenous. The distribution F is then derived endogenously.) Note that for the headcount measure ($\alpha = 0$) $V = F(z)$: the vulnerability measure reduces to the probability that the household will experience poverty (in the sense that $c < z$). Hence probability measures of vulnerability (used, for example, by the World Bank) can be seen as special cases of (1).[3]

To apply (1) one needs an estimate of the distribution F. There are several approaches. First, one can use cross-section survey data to estimate the distribution of consumption (at a point in time) *across households* and use this (for each household) as F, *i.e.* as the distribution of consumption *across states of nature*. This would be valid if the distribution of consumption had converged to the ergodic distribution and if this distribution was the same for all households. The homogeneity assumption that observed consumption represents draws from a single distribution can be relaxed by controlling for observable household characteristics (Kamanou and Morduch, 2001) such as location or educational attainment, but this shifts the problem to a lower aggregation level.

Secondly, if panel data are available then F can be estimated as the distribution of consumption *across time*, for a particular household. In this case the intertemporal mean $\bar{c} = (c_1 + \cdots + c_T)/T$ is considered as the permanent component of consumption and all deviations from this mean as transient. Jalan and Ravallion (2000) use this method for China and McCulloch and Baulch (2000) do so for Pakistan. This method allows for inter-household heterogeneity but, as before, imposes the assumption that F is stationary. When in fact the deterministic part of consumption follows a negative trend this methodology will underestimate vulnerability by treating low consumption levels as unlikely deviations from the intertemporal

---

[3]The World Bank defines vulnerability as "the risk today to fall below the poverty line tomorrow" (Coudouel *et al.*, 2001, p. 37).

mean.[4] Conversely, when there is a positive trend (as in the Zimbabwe data set analysed by Elbers *et al.* (2003) which exhibits very rapid growth or in Scott's (2000) analysis of Chilean data for the period 1968-86) then vulnerability would, of course, be *over*estimated.[5]

A third method is to regress changes in consumption on household characteristics using bootstrapping to generate a distribution of shocks from the regression residuals (Kamanou and Morduch, 2001). The estimated equation can then be used to predict future consumption and vulnerability can be calculated by using the distribution around this mean.

The expected poverty concept is unattractive. For example, one would want an increase in risk (in the sense of a mean-preserving spread) to *increase* expected poverty (consistent with the reduction in welfare experienced by a risk-averse household) but in the case of the Foster-Greer-Thorbecke class of poverty measures this is guaranteed only for $\alpha > 1$ (Ligon and Schechter, 2003; cf. the earlier results of Ravallion, 1988). This rules out the two most popular members of the class: the poverty gap measure ($\alpha = 1$) would record no change when risk increases whereas the headcount ($\alpha = 0$) would (perversely) show an *improvement*, a reduction in expected poverty. Conversely, $\alpha > 1$ implies that the degree of absolute risk aversion increases with consumption (for $c < z$), contrary to much of the evidence available.

Ligon and Schechter (2003) define a utilitarian vulnerability measure. The vulnerability of household h is given by

$$V_h = u_h(z) - E u_h(c_h)$$

where u_h is a weakly concave function (possibly, but not necessarily the household's utility

---

[4]This would affect the results of McCulloch and Baulch: their data have a negative trend.

[5]Ravallion (1988) does not measure vulnerability but he considers a closely related question: the welfare cost of variability. His money-metric for this is the amount by which income - when stabilised at the intertemporal mean - would have to be reduced for poverty to be equal to its intertemporal mean. Clearly, this procedure is very similar to the second method.

function) and z can be seen as a certainty equivalent poverty line. For $V_h > 0$ the household would be considered vulnerable.[6] The measure can be decomposed as

$$V_h = [u_h(z) - u_h(Ec_h)] + [u_h(Ec_h) - Eu_h(c_h)]$$

where Ligon and Schechter interpret the first term as the non-random part of vulnerability ('poverty') and the second term as the effect of risk. In applying this decomposition the unconditional expectation Ec_h is calculated as the mean of monthly consumption data over a period of a year.

In our view methods for estimating vulnerability on the basis of equation (1) are unsatisfactory for four reasons. First, they rely on strong statistical assumptions, *e.g.* homogeneity of the distribution of consumption. In the absence of panel data this is, of course, to some extent unavoidable. However, we find that even with cross-section data some of the non-stationarity can be captured by appropriate modelling.

Secondly, the methods are essentially static: they focus on a household's welfare at a particular moment. This makes sense only if $F(c)$ is the ergodic distribution of household consumption.

Thirdly, the methods ignore any behavioral response to risk. (This applioes not only to the expected poverty measures but also to the Ligon and Schechter approach.) Implicitly they assume that risk affects the volatility of consumption around the mean Ec_h but not the

---

[6]Equivalently, vulnerability could be measured as $V_h = \max(u_h(z) - Eu_h(c_h), 0)$ in analogy with a Foster-Greer-Thorbecke $\alpha = 1$ poverty measure. Note that the cutoff at z is applied to expected utility rather than to consumption c_h in a particular state of nature. The alternative $V_h = u_h(z) - Eu_h(\min(c_h, z))$ is problematic when amended so as to apply in a dynamic context. While in a static context the evaluator may consider outcomes above a certain level irrelevant for the assessment of vulnerability, this position is difficult to maintain in a dynamic setting where a household's consumption at time t affects its scope for avoiding disastrous outcomes at future dates.

mean itself. In a structural model this is a most unlikely outcome. In general risk affects the savings and investment decisions of households and thereby the mean consumption level. This effect can be very large but can be measured only if a structural dynamic model is used.[7] In the absence of such a model one will underestimate risk reducing mechanisms. For example, the Ligon-Schechter decomposition will overestimate the poverty component of vulnerability and underestimate the effect of risk by treating mean consumption as independent of risk. A household which had adopted complete self-insurance (at the cost of lower mean consumption) would be (erroneously) considered as unaffected by risk.

Finally, a household's vulnerability can be low either because it is not exposed to large shocks or because it is able to cope effectively with shocks. Policy makers would want to distinguish between the two cases. For example, a household may achieve consumption smoothing through means which are unnecessarily costly in terms of growth. There would then be a case for intervention (providing insurance to substitute for consumption smoothing through (dis)saving of liquid assets) but the case can be identified only if the household is classified as vulnerable in spite of its consumption smoothing. If this distinction is to be made we must be able to estimate both actual and counterfactual vulnerability, the latter for the hypothetical case where the household faces no shocks. This requires, again, a structural model so that behavioral responses to risk are taken into account. With a proper structural model vulnerability can be assessed separately with and without risk so that the cost of the household's coping mechanism can be estimated.

---

[7]Thorbecke (2003, p. 13) notes that a fundamental feature of the approach we adopt in this paper is "that it incorporates the possibility of households deciding within an intertemporal framework to reduce their mean consumption to reduce consumption variability and risk".

3 Simulation-based Vulnerability Estimates

Our starting point is a Ramsey model: there is a single good, used for consumption, as a store of value and as a productive asset and agents optimize over an infinite horizon.[8] At time $t = 0$ household h solves[9]:

$$W_0 = W(w(k_{h0}, s_{h0}^y, s_{h0}^k)) = \max_{\{c_{ht}, k_{h,t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_{ht}) \tag{2}$$

subject to

$$c_{ht} = w_{ht} - k_{h,t+1}$$

$$w_{ht} = s_{ht}^y a_{ht} f_h(k_{ht}) + s_{ht}^k (1-\delta) k_{ht}$$

$$\text{for } t = 0, 1, 2, .. \text{ and } k_{h0}, s_{h0}^y, s_{h0}^k \text{ given}$$

where c denotes consumption, k the capital stock, w wealth on hand, u the instantaneous utility function, β a discount factor ($0 < \beta < 1$), and δ the depreciation rate. Households are indexed by h and t denotes time. Households have identical preferences but they differ in total factor productivity (measured by the parameter a_{ht}), in exposure to shocks and in assets (k_{ht}).

Unlike in the original Ramsey model, the household is exposed to risk: income $af(k)$ and assets $(1-\delta)k$ are both affected by shocks: s^y, s^k where $E \ln s^y = E \ln s^k = 0$. These income and asset shocks have both idiosyncratic and covariant components:

$$s_{ht}^y = (\varepsilon_t^r)^\pi \varepsilon_{ht}^y$$

$$s_{ht}^k = (\varepsilon_t^r)^\pi \varepsilon_{ht}^k.$$

---

[8] This section draws on Elbers *et al.* (2003) where the model and estimation method are described in more detail.

[9] The extension to W_t for $t = 1, 2, ..$ is straightforward.

9

We identify the covariant shocks ε^r with rainfall (denoted by the superscript r) and measure their importance by the elasticity π. We assume that the distributions of $\varepsilon_{ht} = (\varepsilon_{ht}^y, \varepsilon_{ht}^k)$ and ε_t^r are lognormal, independent of each other and across time and that $\ln \varepsilon_{ht}$ has correlation matrix

$$\begin{pmatrix} a_1^2 & a_1 b_1 \\ a_1 b_1 & b_1^2 + b_2^2 \end{pmatrix}.$$

When the household decides on c_t and k_{t+1} both k_t and the realizations (s_{ht}^y, s_{ht}^k) are known. Future shocks are, of course, unknown but the household does know the distributions of these shocks.[10] Note that wealth on hand, w, is a function of the capital shock and the current shocks:

$$w_{ht} = w(k_{ht}, s_{ht}^y, s_{ht}^k).$$

Solving the model involves finding a policy function

$$\varphi(w(k, s^y, s^k)) = \arg\max_{\tilde{k}} u(w(k, s^y, s^k) - \tilde{k}) + \beta EW(w(\tilde{k}, \tilde{s}^y, \tilde{s}^k))$$

where k and $\tilde{k}$ denote the capital stock at the beginning and the end of the current period and similarly s and $\tilde{s}$ denote current and future shocks. The policy function φ maps the current (k, s^y, s^k) into $\tilde{k}$, next period's k. Hence φ can be seen as an investment function, giving k_{t+1} as a function of wealth on hand w_t (itself a function of the capital stock k_t) and the current shocks s_t^y and s_t^k.

The Ramsey model serves only as an example. It fits the Zimbabwe data quite well but it may not be appropriate in other contexts.

In this stochastic Ramsey model risk affects household behaviour in two ways. First, if the household perceives a change in the distribution of shocks (*e.g.* an increase in rainfall risk in

---

[10]This is a strong assumption: in reality these distributions may change over time and it will take time for agents to learn about such changes.

the form of a mean preserving spread of the covariant shock ε_t^r) then it will, in general, adjust its policy function φ so that (for the same values of the capital stock k_t and the shocks s_t^y, s_t^k) it will choose different values of k_{t+1} (and hence c_t). This effect of a change in risk on the household's policy function we term the *ex ante* effect. There also is an *ex post* effect: the change in risk will affect the size of the realised shocks so that the optimal values of k_{t+1} and c_t (controlling for k_t) are affected, even for an unchanged policy function.

Elbers *et al.* (2003) estimate this model using a simulated pseudo maximum likelihood technique on an 18-year panel data set for 158 smallholder households in Zimbabwe. The Zimbabwean households were exposed to massive shocks in the period 1982-2000, including a very serious drought in 1991/2. In this data set there are observations on livestock holdings and we identify this variable with the capital stock k.

Elbers *et al.* assumed that $u(c)$ is a CRRA function: $u(c) = c^\gamma$, that the production function is of the CES type:

$$f(k) = (1 + \psi(k^{-\rho} - 1))^{-1/\rho}$$

and that total factor productivity is a function of the household's size (hhsize) and the highest educational attainment of its adult members (ed):

$$a_{ht} = (\alpha_0 + \alpha_1 \text{hhsize} + \alpha_2 \text{ed}).$$

We use the Elbers *et al.* parameter estimates, summarized in Table 1. These estimates imply a fairly high elasticity of factor substitution in agricultural production: since ρ is close to -0.5, the elasticity is close to 2. They also imply a close to unitary degree of relative risk aversion so that the utility function is approximately $u(c) = \ln c$.

Table 1: Model Parameters

parameter	estimate
α_1/α_0	-.006887
α_2/α_0	.037815
$_2$	54.038
ψ	0.5315
ρ	-0.5394
γ	0.0082
β	0.7490
δ	0.1330
π	0.0330
a_1	0.2691
b_1	0.2394
b_2	0.1389

Source: Elbers *et al.* (2003)

W measures the household's perceived welfare, taking into account the risks it faces and the scope for consumption smoothing through changes in assets. We estimate vulnerability as the shortfall of W from the welfare level the household would attain if its consumption would be equal to the poverty line in every period: $c_t = z$ for $t = 0, 1, 2, ..$:

$$V_t = W(z) - W_t \tag{3}$$

12

where

$$W(z) = \sum_{t=0}^{\infty} \beta^t u(z) = \frac{u(z)}{1 - \beta}. \tag{4}$$

Note that this measure does not suffer from the problems identified in the previous section: there is no need to assume homogeneity or stationarity of the distribution of consumption (indeed we need not make any assumption: the distribution is determined endogenously); since V is a linear transformation of the welfare measure W the vulnerability measure cannot be inconsistent with household welfare; the measure is not static but is based on utility over an infinite horizon; and, finally, the impact on V of the *ex ante* and *ex post* effects of risk can easily be identified. This involves solving the model and calculating V (a) under the assumption that there is no risk; (b) under the assumption that the household correctly perceives the distribution of the shocks it faces but experiences no shocks ($s^y = s^k = 1$ throughout); and (c) under the assumption that the household experiences shocks drawn from the (correctly perceived) distributions.

Figure 1 illustrates the effect of risk on asset accumulation in this model. The Figure shows for one of the households in the Zimbabwe sample how the optimal value of the capital stock (K/L_e: livestock per labour in efficiency units) evolves over a 50-year period. The "sample path under risk" represents for a selected household the path of k_{ht} for a particular (randomly drawn) series of shocks. Note that the shocks are very large: for much of the period asset ownerships changes by 50% in one or two years. The "no risk" path shows that in the absence of risk the household's capital stock would have grown steadily and quite rapidly, from an initial value of about 0.5 to a steady state value of about 2.7. The "no *ex post* risk" path shows that the *ex ante* effect of risk reduces this steady state value dramatically: to about 1.7. When both the *ex ante* and *ex post* effects are taken into account the path depends, of course, on the realized shocks s^y, s^k. The "average under risk" path is calculated as the mean over

13

100,000 sample paths. This path shows that the *ex post* effect further reduces growth over the simulation period.

Note that the effect of risk on capital accumulation is very strong and that about two-thirds of the total effect is accounted for by the *ex ante* effect of risk. Vulnerability methods which treat the mean over time of a household's consumption as the riskless counterfactual (*e.g.* Ravallion, 1988; Ligon and Schechter, 2003) would in this case miss most of the story: they would erroneously treat the "average under risk" long run value (about 1.5) as the riskless counterfactual, the deterministic long-run value (of about 2.7). If measures of chronic poverty are based on mean consumption over time then a large part of chronic poverty could in fact reflect risk.

The large effect of risk on asset accumulation had already been identified in Elbers *et al.* (2003). We now consider the implications of the stochastic Ramsey model for the measurement of vulnerability. We do this by considering the estimated model as the data generating mechanism: the model is used to generate household-specific paths of consumption or asset holdings. We then ask whether existing vulnerability measures would correctly identify the most vulnerable households.

Figure 2 illustrates the difficulty of doing so. The Figure shows for five different combinations of the initial capital stock and total factor productivity (k_{h0}, a_h) how welfare (W) - calculated as "the average under risk" path in Figure 1, *i.e.* as the mean over a large number of simulated paths - changes over time. Clearly, in the short run differences between households in welfare (and hence in vulnerability in the way we define it) reflect differences not only in k_{h0}, a_h but also in shocks. Here we have isolated the effects of differences in k_{h0}, a_h by averaging the value of W over 100,000 paths. Of the remaining two sources of heterogeneity, differences in initial assets matter only in the short run. A household's welfare in the very long run is independent of k_{h0}: it depends only on a_h. The effect of this can be dramatic, as illustrated by
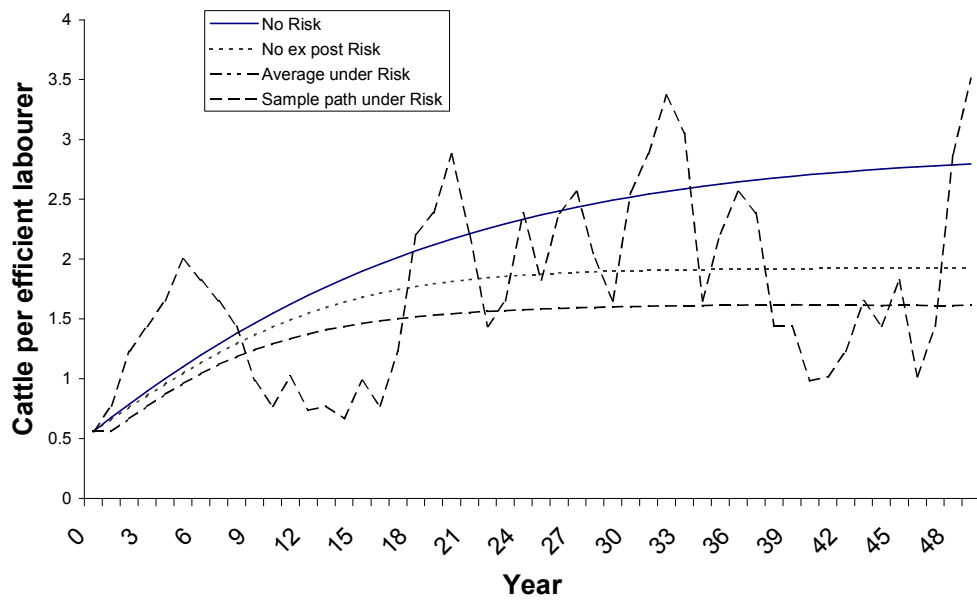
14

Figure 1: Asset Accumulation and Risk (selected household). Source: Elbers *et al.* (2003).

Welfare flow

Sample K0-TFP combinations
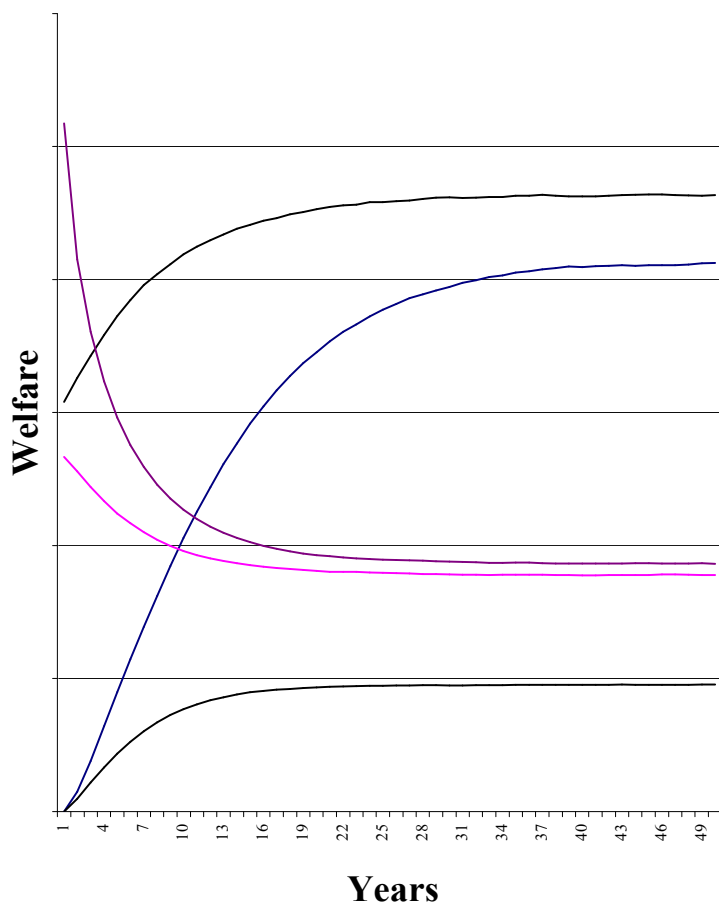

Figure 2: Welfare (W) for selected combinations of total factor productivity and initial capital.

the two paths in Figure 2 which initially have the lowest W value. One of these households is very much more productive (an advantage which is initially offset by a lower value of k_{h0}) and eventually reaches a very much higher welfare level.

This illustrates that whether or not a household is to be considered as vulnerable depends on the time period considered. The Figure makes this point in a striking way: many of the path actually cross each other. This implies that the ranking of households (in terms of welfare or vulnerability) changes over time. For example, the household which in Figure 2 initially has the highest value of W ends up in third place: apparently its relatively low productivity was offset by a high initial value of k_{h0}.

4 Comparing Vulnerability Measures

In practice the data needed to estimate a structural dynamic model such as the stochastic Ramsey model will often not be available. How far can we then get with the less data intensive regression-based estimates discussed in section 2?

Theory - as exemplified by the stochastic Ramsey model of the previous section - implies a mapping from assets (k), shocks (s), productivity determinants (x), and characteristics of the distribution of shocks (σ) to consumption (c):

$$c_{ht} = \xi(k_{ht}, s_{ht}, x_{ht}, \sigma_{ht}) \tag{5}$$

and from this mapping one can derive appropriate vulnerability measures, such as our measure V, based on the household's expected discounted utility. However, in practice vulnerability measures are not based on equation (5) but typically on regressions of c_{ht} on x_{ht} and possibly also on s_{ht}. This approach is problematic in several ways.

First, functional form restrictions have to be imposed and these are often highly restrictive. For example, without interaction terms a linear specification makes the effect of s on c

independent of x. There is no theoretical justification for such independence.

Secondly, by leaving out assets (k) the regression suffers from omitted variable bias. This could be serious: if two households are observed after being hit by a negative shock and if they are identical in all respects except for the level of assets then their consumption decisions may well be very different: the household with the higher k can better afford to smooth consumption by using its assets.

Thirdly, omitted variable bias is also introduced by the exclusion from the regression of the shock characteristics σ_{ht}. Recall that a change in risk affects household behaviour both *ex ante* (the policy function φ will be affected, *i.e.* the household will for the same values of (k, s^y, s^k) decide on a different level of investment and consumption) and *ex post (*since the shocks (s^y, s^k) are now drawn from different distributions). If s is included in the regression but σ is not, then this *ex post* effect can in principle be estimated but the *ex ante* effect will be missed. This is potentially serious: if one would like to estimate how much a policy-induced reduction in risk would contribute to welfare then one needs to have estimates of both effects.

Finally, implicitly the regression treats the distribution of c across households (conditional on the regressors) as the distribution of shocks. This is appropriate only if households do not change this distribution by changing their asset holding, in other words if the distribution of assets has converged to the long run (ergodic) distribution. (This assumption is certainly not appropriate in our case. For example, mean assets of the rural households in the Zimbabwean sample in 1983 - shortly after they were resettled - were only about 20% of the mean of the ergodic distribution.)

For all these reasons, vulnerability measures based on the usual consumption regressions may be wrong. The question is: how wrong? We investigate this with a series of experiments. We take the estimated stochastic Ramsey model as the correct model and use it as the data generating process for a series of regressions (Table 3). In each case the data generated are for

1981, *i.e.* one year ahead.

The first regression relates consumption only to household-specific total factor productivity.[11] The performance of this regression is very poor. Clearly it cannot serve as a basis for identifying vulnerable households. The second regression includes shocks experienced as additional regressors.[12] This results in a substantial improvement of the fit. Finally, in the last regression we also include the household's asset position (in terms of cattle). This leads to an excellent fit: $R^2 = .95$, in spite of the fact that the equation is *ad hoc* from a theoretical point of view.

---

[11] In practice total factor productivity (tfp) is not observed and proxied by household characteristics. In our model tfp is determined by household size and education. To avoid artificial errors from choosing the wrong functional form in the regression we have used tfp as regressor.

[12] In practice the researcher would have at best partial information on these shocks.

Table 2

Consumption regressions on simulated data

	TFP only		TFP and shocks		TFP, cattle, shocks	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Constant	-19.64	17.5	-20.84	15.6	-4.61	0.501
TFP	8.474	7.22	7.026	6.42	0.522	0.094
TFP^2	-0.760	0.737	-0.627	0.656		
Cattle					3.485	0.077
s^y			4.076	1.402	3.122	0.373
s^k			0.755	1.305	0.262	0.347
R^2	0.050		0.258		0.948	
Dependent variable is simulated after-shock 1981 consumption						

We now consider how such regressions can be used to identify vulnerable households. Figure 3 shows how the sample households are initially distributed over the (cattle,TFP) plane. The contour lines show combinations of productivity and cattle for which household welfare (W) and hence vulnerability (V) are the same. Note that the indifference curves are very steep: V is very sensitive to changes in initial asset ownership. This reflects the short horizon of one year used in the present exercise. If we take a longer horizon (say 5 years) the indifference curves would become flatter. In the limit, when initial conditions are no longer relevant, the curves are horizontal. (Figure 2 makes the same point.)

If instead of the "poverty gap" ($\alpha = 1$) vulnerability concept of equation (3), we use a "headcount" concept then measuring vulnerability simply amounts to identifying those house-

holds with $W < W(z)$. If we were to use the official Zimbabwe poverty line then all households in our sample would be classified as vulnerable. For our present purpose of assessing how well existing vulnerability measures succeed in identifying the vulnerable, such a high value of z is not useful. We therefore choose a lower value of z, namely the value which classifies half of the sample as poor. In Figure 3 the line with large dashes separates the 50% worst off from from 50% best off: dots below this line represent households which are vulnerable at the chosen (obviously arbitrary) poverty line.

If vulnerability was measured on the basis of the first regression, *i.e.* on the basis of total factor productivity, and if in addition the researcher knew that 50% of the population was vulnerable then he would identify vulnerable households as those below the bold horizontal line in Figure 3. Clearly many vulnerable households would not be identified as such and vice versa. This is shown in Table 3 which shows that almost 40% of the households are misclassified by this regression. (The error would be larger if, realistically, the researcher did not know how many households were poor and if he could observe some of the household characteristics which determine productivity but not productivity itself.) This suggests that regression-based estimates which measure vulnerability on the basis of observable proxies for total factor productivity may be very inaccurate.

The two steeper lines in Figure 3 correspond to regressions with cattle ownership as the only explanatory variable (solid line) or with cattle ownership and shocks as the regressors (short dashes). Very few households are misclassified by these regressions. (The fact that the indifference curve corresponding to these two regression are steeper than the true indifference curve reflects the fact that the Ramsey program also takes into account the long-run effects of total factor productivity.) Indeed, as shown in Table 3, in either case only two households are misclassified. This suggests that in vulnerability analysis the perfect need not be the enemy of the good. When the panel data needed to implement a structural model such as the stochastic
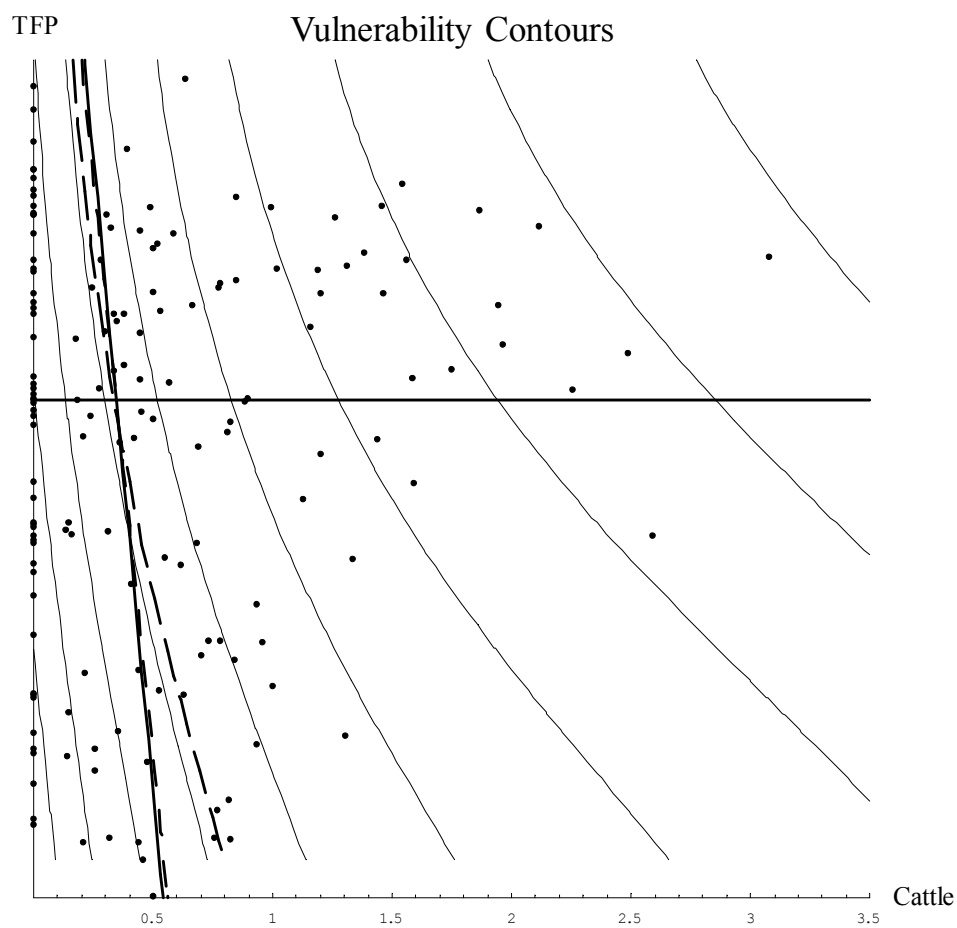
21

Figure 3: Classification of households as vulnerable by various methods.

Ramsey model are not available, then it is still possible to achieve a very good approximation on the basis of cross-section data, provided asset data are included in the regression.

Table 3		
Classifying the 50% poorest households.		
Stochastic Ramsey vs. regression models		
Model	TFP only	TFP, cattle
		and shocks
Correct poor	48	78
Correct non-poor	48	78
False poor	31	1
False non-poor	31	1
Total	158	158

5 Conclusion

It has long been recognised that poverty measures based on cross-section data may be misleading indicators of household welfare if there is substantial unobserved heterogeneity in the sample, if they are observed when the distribution of assets differs substantially from the ergodic distribution, and, perhaps most importantly, if they face risk. Existing vulnerability measures try to incorporate the effect of risk on welfare. However, most measures currently used lack a theoretical basis while the measures which do have a theoretical basis are derived in a static framework. We have argued that vulnerability can be measured rigorously as the shortfall from some critical value of the welfare of a household which solves an intertemporal optimisation model under risk. Using such a model (a stochastic Ramsey model estimated on

panel data for smallholders in Zimbabwe) we showed that failing to distinguish between the *ex ante* and *ex post* effects of risk may lead to large errors in estimates of chronic and transient poverty.

Our analysis makes clear that vulnerability depends on the time horizon considered. In particular, if one takes a longer term perspective, vulnerability is less sensitive to initial conditions and, conversely, more sensitive to permanent productivity differentials between households.

We used the estimated model to assess the accuracy of existing methods in identifying vulnerable households. Regression-based methods using proxies for total factor productivity (*e.g.* education or soil quality) or measures of shocks experienced by the household as the only regressors can be very misleading. Our results show that a vast improvement can be obtained by including asset measures in the regression. Using the estimated Ramsey model as a data generating laboratory we found that simple, regression-based methods can accurately classify households provided asset data are included as regressors.

References

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

Christiaensen, Luc J. and Kalanidhi Subbarao (2001), 'Towards and Understanding of Vulnerability in Rural Kenya', mimeo, World Bank.

Coudouel, A., J. Hentschel and Q. Wodon (2001), 'Well-Being Measurement and Analysis', mimeo, World Bank.

Deaton, Angus (1990), 'Saving in Developing Countries: Theory and Review', *World Bank Economic Review, Supplement: Proceedings of the 1989 ABCDE*, pp. 61-96.

Dercon, Stefan (2002), 'Risk, Insurance and Poverty: a Review', forthcoming in Stefan

Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER.

Dercon, Stefan and Pramila Krishnan (2000), 'Vulnerability, Seasonality and Poverty in Rural Ethiopia', *Journal of Development Studies*, vol. 36, pp. 25-51.

Dercon, Stefan and Pramila Krishnan (2003), 'Risk Sharing and Public Transfers', *Economic Journal*, vol. 113, pp. C86-C94.

Elbers, Chris, Jan Willem Gunning and Bill Kinsey (2003), 'Growth and Risk: Methodology and Micro Evidence', Tinbergen Institute Discussion Paper 03-68/2.

Jalan, Jyotsna and Martin Ravallion (1998), 'Transient Poverty in Post-Reform Rural China', *Journal of Comparative Economics*, vol. 26, pp. 338-357.

Jalan, Jyotsna and Martin Ravallion (2000), 'Is Transient Poverty Different? Evidence for Rural China', *Journal of Development Studies*, vol. 36, pp. 82-99.

Kamanou, Gisele and Jonathan Morduch (2001), 'Measuring Vulnerability to Poverty', mimeo.

Ligon, Ethan and Laura Schechter (2003), 'Measuring Vulnerability', *Economic Journal*, vol. 13, pp. C95-C102.

McCulloch, Neil and Bob Baulch (2000), 'Simulating the Impact of Policy upon Chronic and Transient Poverty in Rural Pakistan', *Journal of Development Studies*, vol. 36, pp. 100-130.

Morduch, Jonathan (1994), 'Poverty and Vulnerability', *American Economic Review, Papers and Proceedings*, vol. 84, pp. 221-225.

Pritchett, Lant, Asep Suryahadi and Sudarno Sumarto (2000), 'Quantifying Vulnerability to Poverty: a Proposed Measure, Applied to Indonesia', mimeo, Social Monitoring and Early Response Unit (SMERU), Jakarta.

Ravallion, Martin (1988), 'Expected Poverty under Risk-Induced Welfare Variability', *Economic Journal*, vol. 98, pp. 1171-1182.

Scott, Christopher D. (2000), 'Mixed Fortunes: a Study of Poverty Mobility among Small

Farm Households in Chile, 1968-86', *Journal of Development Studies*, vol. 36, pp. 155-180.

Thorbecke, Erik (2003), 'Conceptual and Measurement Issues in Poverty Analysis', paper presented at the WIDER Conference on "Inequality, Poverty and Human Well-Being", Helsinki, Finland, May 30-31.

World Bank (2001), *World Development Report 2000/2001: Attacking Poverty*, Washington, DC: World Bank.