Vulnerability as a Measure of Chronic Poverty

Laura Schechter

October 5, 2006

Economists have traditionally used static poverty measures to estimate well-being, target aid, and determine progress towards development goals. This is useful as a starting point, and gives a good snapshot of the household's situation at one point in time, but in the end, what we care about is the standard of living of the household over its entire lifetime. A household's well being depends not just on its expenditures at one point in time, but on its average expenditures over its lifetime. Because of this, researchers now want to measure chronic poverty, in addition to static poverty. Because this is a newer concept, the literature has not yet converged on one measure which is acceptable to all, as, arguably, the literature on static poverty has.

We argue that using a measure of chronic poverty based on the average level of expenditures still misses an important portion of economic well-being. For *Voices of the Poor*, the 2000/01 World Development Report, the World Bank interviewed tens of thousands of poor people in countries across the globe. These poor people stated that part of what makes poverty so unbearable is the instability and unpredictability they face, leading to a feeling of vulnerability. Thus, we propose a measure of chronic poverty, which we call vulnerability, which incorporates the risk households face as well as their average level of expenditures.

The majority of vulnerability measures proposed thus far are static measures. (Elbers & Gunning (2006) is one exception.) Thus they are backward-looking rather than forward-looking. In this paper we begin with a static measure of vulnerability and analyze the South African KIDS data using this measure. Then, we propose four different ways in which one could make the measure more forward looking. Lastly, we choose one of these methods and apply it to the South African data.

1 Measurement of Vulnerability

Note that in order to measure chronic vulnerability there are two main steps involved. The first is to decide what measure(s) one wants to use. One can use measures based on expected utility as suggested here, measures based on expected poverty, asset-based measures, or any other measure. The second step is to estimate the distribution of expenditures (or assets) which a household may face. A preliminary discussion of this second step can be found in Ligon & Schechter (2004), with suggestions for how to progress.

In this paper we use the measure of welfare suggested by Ligon & Schechter (2003). They suggest a measure of household (or individual) welfare which they term vulnerability, and this section of the paper borrows heavily from their set-up. This measure takes into account both poverty and risk. Given some utility function U, they define the vulnerability, V, of the household by the function

$$V = U(z) - \mathrm{E}U(c_i). \tag{1}$$

Here c_i is household-level consumption and z is some certainty-equivalent consumption (perhaps the poverty line) such that if household i had certain consumption greater than or equal to this number, we would not regard the household as vulnerable. Note that one can decompose this measure of vulnerability into two distinct components.

$$V = [U(z) - U(\mathrm{E}c^i)] + [U(\mathrm{E}c^i) - \mathrm{E}U(c^i)]. \tag{2}$$

The first term in brackets measures chronic poverty. This is not the household's poverty at one moment in time, but its expected level of poverty over time. The second term in brackets is a measure of the risk faced by the household. Two household may both be chronically poor, but the household which faces more variance in its expenditure pattern should probably be considered worse-off.

This risk measure can be further decomposed into two distinct measures of risk, one aggregate and the other idiosyncratic. Let $\mathrm{E}(c^i|\bar{x}_t)$ denote the expected value of consumption, c^i, conditional on a vector of aggregate variables $\bar{x}_t$. Then we can rewrite vulnerability as

$$\begin{aligned} V = \quad & [U(z) - U(\mathrm{E}c^i)] && \text{(Poverty)} \\ + \quad & [U(\mathrm{E}c^i) - \mathrm{E}U(\mathrm{E}(c^i|\bar{x}_t))] && \text{(Aggregate risk)} \\ + \quad & [\mathrm{E}U(\mathrm{E}(c^i|\bar{x}_t)) - \mathrm{E}U(c^i)]. && \text{(Idiosyncratic risk)} \end{aligned}$$

Here the second term expresses the aggregate risk facing the household, while the third filters out the aggregate component of risk to leave only the component of idiosyncratic risk.

In the presence of measurement error, using observed consumption to measure vulnerability would lead the analyst to confute measurement error with idiosyncratic risk. To avoid this problem, we further decompose our measure of idiosyncratic risk into risk which can be attributed to variation in k observed time-varying household characteristics $x_t^i = (x_{1t}^i, \ldots, x_{kt}^i)$ and a risk which can neither be explained by these characteristics, nor aggregate variables, but which is due instead to variation in unobservables and to measurement error in consumption. Thus, rewriting the expression for vulnerability yields

$$\begin{aligned}
V = \quad & [U(z) - \mathrm{E}U(\mathrm{E}c_t^i)] \qquad \text{Poverty} \\
+ \quad & [U(\mathrm{E}c_t^i) - \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t))] \qquad \text{(Aggregate risk)} \\
+ \quad & [\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t)) - \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_t^i))] \qquad \text{(Idiosyncratic risk)} \\
+ \quad & [\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t x_t^i)) - \mathrm{E}U(c_t^i)] \qquad \text{(Unexplained risk \& measurement error).}
\end{aligned}$$

We can further decompose "Explained idiosyncratic risk" into k distinct sources. If the k variables x_{jt}^i are not mutually orthogonal they can first be orthogonalized via a Gram-Schmidt procedure. If they are mutually orthogonal, then we simply write explained idiosyncratic risk as

$$\begin{aligned}
\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t)) \quad - \quad & \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_t^i))] = [\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t)) - \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_{1t}^i))] \\
+ \quad & [\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_{1t}^i)) - \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_{1t}^i, x_{2t}^i))] \\
\vdots \quad & \\
+ \quad & [\mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_{1t}^i, \ldots, x_{(k-1)t}^i)) - \mathrm{E}U(\mathrm{E}(c_t^i|\bar{x}_t, x_{1t}^i, \ldots, x_{kt}^i))].
\end{aligned}$$

Suppose, for example, that we have data on household assets, household income, and the number of days someone in the household was sick. We would denote by x_{1t}^i the part of household assets which is orthogonal to household and time effects; by x_{2t}^i the part of household income in household i orthogonal to household effects, time effects, and household assets, and by x_{3t}^i the part of days of sickness in household i orthogonal to all the other variables. Thus, using our example, the first bracketed term of (3) provides a measure of the welfare loss which can be predicted using variation in household is assets; the second bracketed term the change in prediction if we include data on income, and so on. Notice that, if two idiosyncratic risk variables are highly correlated, the one which is included first in the decomposition will pick up most of the effects on idiosyncratic risk.

2 Estimation of Vulnerability Using Households' Past Experiences

Two additional steps are required before one can actually use data to compute a household's vulnerability. First, one must choose the function U. Second, one must devise a way to estimate the conditional expectations which figure in our vulnerability measure. Here, we assume that $U(c) = (c^{1-\gamma} - 1)/(1-\gamma)$ and that $\gamma = 2$. The original paper by Ligon and Schechter assumes a stationary environment and optimally predicts c_t^i in a least-squares sense.

$$\ln c_t^i = \alpha^i + \eta_t + x_t^i \beta + v_t^i \tag{3}$$

Here α^i captures the influence of fixed household characteristics on predicted household consumption while η_t captures the influence of aggregate risk. The shock variables x_t^i have been orthogonalized via the Gram-Schmidt procedure.

2.1 Application to South Africa

This section carries out the above vulnerability analysis using the KIDS data from South Africa in 1993, 1998, and 2004. For more information on the data see May et al. (2006). The poverty line determined by Alderman et al. (2000) is 322 rand per month in 2000 prices. Table 1 uses data from South Africa to demonstrate vulnerability and its breakdown into poverty and risk, as well as the correlates of all three.

Carter & Maluccio (2003) find, using the same data set, that children whose households experienced shocks during the first three years of their lives are more likely to be stunted, especially if their neighbors experienced losses at the same time. This suggests that both idiosyncratic and aggregate risk could be important in this situation. This is corroborated by the evidence in Table 1 which shows that risk accounts for approximately 75% of vulnerability, while poverty only accounts for 25%. This contribution of risk to vulnerability is higher than has been found in other data sets (46% in

Table 1: Determinants of Vulnerability

	Vulnerability =	Poverty +	Risk
	0.2033*** =	0.0499** +	0.1534***
	(1)	(2)	(3)
Hh Size in 93	.078***	.071***	.007***
	(.005)	(.005)	(.001)
Assets in 93	-3.65e-06**	-3.12e-06**	-5.28e-07**
	(1.70e-06)	(1.50e-06)	(2.59e-07)
Unearned Inc in 93	-.0001***	-.0001***	-1.00e-05
	(.00004)	(.00004)	(9.60e-06)
Educ Labor in 93	-.135***	-.121***	-.014***
	(.010)	(.009)	(.002)
Indian	-.540***	-.460***	-.080***
	(.030)	(.026)	(.007)
Pop Density > 500	-.263***	-.234***	-.028***
	(.035)	(.031)	(.007)
Const.	.109**	-.046	.155***
	(.050)	(.043)	(.011)
Obs.	715	715	715
R^2	.536	.539	.293

OLS, bootstrapped standard errors.
*-90%, **-95%, and ***-99% significant.

monthly data from Bulgaria and 51% in yearly data from Vietnam (Ligon & Schechter 2004)). This suggests that households in South Africa face a relatively large amount of risk, and that looking simply at measures of chronic poverty, rather than chronic vulnerability, would ignore a large part of what effects household welfare.

Table 1 goes on to examine which variables are correlated with vulnerability, poverty, and risk. To do this, we regress vulnerability, poverty, and risk, on a set of explanatory variables. The correlates of all three dependent variables are quite similar. Smaller Indian households in rural areas with more physical assets and education are less vulnerable, poor, and risky. When looking at the relative sizes of the coefficients, we find that being Indian and owning assets is more important in the risk regression, while household size,

Table 2: Determinants of Risk

	Risk = 0.1534*** =	Agg Risk + 0.0063*** +	Idio Risk+ 0.0330*** +	Unexp Risk 0.1140***
	(1)	(2)	(3)	(4)
Hh Size in 93	.007***	.0004***	.003***	.003***
	(.001)	(.0001)	(.0006)	(.001)
Assets in 93	-5.28e-07**	-1.64e-08*	-1.37e-07*	-3.73e-07
	(2.59e-07)	(8.60e-09)	(7.50e-08)	(2.31e-07)
Unearned Inc in 93	-1.50e-05	-5.85e-07**	-2.46e-07	-.00002**
	(9.60e-06)	(2.5e-07)	(3.90e-06)	(7.80e-06)
Educ Labor in 93	-.014***	-.0006***	-.003***	-.011***
	(.002)	(.0002)	(.0007)	(.002)
Indian	-.080***	-.003***	-.016***	-.062***
	(.007)	(.0008)	(.003)	(.006)
Pop Density > 500	-.028***	-.001***	-.007***	-.019***
	(.007)	(.0003)	(.002)	(.006)
Const.	.155***	.006***	.022***	.129***
	(.011)	(.001)	(.005)	(.010)
Obs.	715	715	701	701
R^2	.293	.606	.271	.197

OLS, bootstrapped standard errors.
*-90%, **-95%, and ***-99% significant.

education, and population density (urban/rural) are more important in the poverty regression.

Table 2 further breaks down risk into its three components. The first piece is aggregate risk, which is due to changes in consumption which effect everyone in the sample. The second piece is idiosyncratic risk, which is how shocks which we can measure affect vulnerability vis-a-vis consumption. Lastly is unexplained risk which is due partly to shocks not included and partly to measurement error.

Table 2 shows that the lion's share of risk is due to unexplained risk. Unexplained risk accounts for 74% of total risk and 56% of vulnerability. Compare this with Bulgaria where it accounts for 85% of total risk and 39% of vulnerability and Vietnam where it accounts for 79% of total risk and 40%

of vulnerability. The fact that unexplained risk accounts for such a large share of risk and vulnerability in multiple data sets suggests that either the data sets contain a lot of measurement error, or that we, as econometricians, are not doing a good job of explaining and predicting the true risk that households face. We will explore this more in Section 4.

The correlates of each piece of risk are quite similar to those in the previous table. We note that unearned income, educated labor, being Indian, and living in an urban area have relatively large effects on decreasing unexplained risk. This suggests that this unexplained risk is not totally due to measurement error.

Table 3 breaks down idiosyncratic risk into six pieces. The six shocks we have included in our analysis are log total household income, days missed by all household members from work in a period of 14 days due to sickness, household size, uneducated labor, total productive capital, and experiencing victimization such as assault, robbery, rape, or kidnapping. Let's take the example of household size risk. A household's size changes over time and this may affect per capita consumption within the household. Here we measure the risk experienced by a household due to such changes.

Risk due to changes in income risk and changes in household size are the largest components of idiosyncratic risk. Given the AIDS epidemic ravishing South Africa the result regarding household size is not surprising. Idiosyncratic risk due to victimization is the only other piece of idiosyncratic risk which is significantly different from zero, echoing the constant complaints of South Africans of criminality being a major concern. Asset risk, sickness risk, and education risk are all insignificant. [*A next step is to move asset risk to be the first component of idiosyncratic risk. As noted in the previous section, the component which is included first will tend to pick up most of the effects of other variables with which it is correlated. Given that Michael Carter is using the same data set to discuss asset-based poverty traps, we would like to give asset risk the benefit of the doubt by including it first.*] Victimization risk is not correlated with any of the explanatory variables, suggesting that experiencing rape, kidnapping, robbery, or assault is a random occurrence. Smaller households, those with more educated labor, Indian households, and those coming from more urban areas experience less idiosyncratic risk. It is interesting to note that productive capital is only just barely significant in predicting risk.

Table 3: Determinants of Idiosyncratic Risk

	Id Risk =	Inc Risk +	Sick Risk +	HhSize Risk +	Unedu Risk +	Asset Risk +	Victim Risk
	0.0330*** =	0.0192***+	3.4e-06 +	0.0129*** +	0.00009 +	0.00008 +	0.0008***
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hh Size in 93	.003***	.0009***	1.21e-06	.002***	-.00008	.00002	.00009
	(.0006)	(.0002)	(.00002)	(.0005)	(.00007)	(.00003)	(.00007)
Assets in 93	-1.37e-07*	-4.86e-08	2.30e-10	-9.95e-08	-3.01e-09	8.80e-09	5.04e-09
	(7.50e-08)	(4.10e-08)	(4.20e-09)	(6.10e-08)	(8.60e-09)	(2.20e-08)	(1.60e-08)
Unearned Inc in 93	-2.46e-07	-5.28e-06***	-3.33e-09	4.53e-06	1.83e-07	7.81e-08	2.37e-07
	(3.90e-06)	(1.50e-06)	(9.90e-08)	(3.40e-06)	(2.90e-07)	(1.70e-07)	(3.20e-07)
Educ Labor in 93	-.003***	-.001**	-3.84e-06	-.002***	.0002	.00002	-9.82e-06
	(.0007)	(.0006)	(.00005)	(.0006)	(.0001)	(.00005)	(.0001)
Indian	-.016***	-.012***	1.35e-06	-.004**	-.00009	-3.45e-06	-.0002
	(.003)	(.002)	(.00009)	(.002)	(.0003)	(.0001)	(.0004)
Pop Density > 500	-.007***	-.003**	-6.93e-06	-.003**	-.0003	-.00007	-.0001
	(.002)	(.002)	(.0001)	(.001)	(.0003)	(.0001)	(.0003)
Const.	.022***	.020***	3.42e-06	.001	.0005	-.00008	.0001
	(.005)	(.003)	(.0001)	(.003)	(.0006)	(.0002)	(.0004)
Obs.	701	701	701	701	701	701	701
R^2	.271	.136	.008	.192	.021	.032	.015

OLS, bootstrapped standard errors.
*-90%, **-95%, and ***-99% significant.

3 Estimation of Vulnerability Using Households' Potential Future Experiences

The above estimation technique leads to a backward-looking measure of vulnerability in the sense that we use information on what happened to a household in the past, not what could happen to a household in the future. This has been a shortcoming in most of the literature on vulnerability thus-far. Most measures of vulnerability have been rather static. This is an extremely difficult issue as, of course, we would expect predicting the future to be difficult. On the other hand, it is quite important. Below I present two relatively simple ways of attacking this issue, and two more difficult ways.

It is useful to keep in mind that, in starting down this path, it becomes much more difficult to control for measurement error. The analysis in the section gave us a measure of vulnerability which was an upper bound on the true level of backward-looking vulnerability. By subtracting unexplained risk from the measure of vulnerability estimated in the previous section, we could estimate a lower bound on vulnerability. In the discussion below, it is no longer obvious how to account for measurement error in a satisfactory manner. But, given that caveat, we will continue on to look at four potential manners in which to make the measure of vulnerability more forward looking.

3.1 Modeling the Evolution of Shocks

Imagine two individuals with equal probabilities of experiencing a robbery. One of them experiences this victimization in one of the survey years while the other does not. For the first individual both poverty and risk will be overestimated while for the second individual they will both be underestimated.[1]

One could estimate the probability that each household will experience any given shock (robbery, death of livestock, sickness, i.e. x_t^i in Equation (3)) given their fixed characteristics. One would also have to use this information to re-estimate α^i to take into account the misestimation of poverty in addition to the misestimation of risk.

To be a little more specific, in the South Africa data set we have three

[1] At the aggregate level this over and underestimation will be less severe, as the sample size increases. Thus, measures of vulnerability for the population as a whole will be less problematic than individual vulnerability rankings.

observations (1993, 1998, and 2004) for each individual regarding the vector of shocks x_t^i. Using the observed data, we can predict the actual distribution of shocks x_t^{*i}. Consider the case with no aggregate risk, in which case we would estimate the equation

$$\ln c_t^i = \alpha^i + x_t^i \hat{\beta} + v_t^i$$

or, equivalently,

$$\ln c_t^i - \overline{\ln c^i} = (x_t^i - \overline{x^i})\hat{\beta} + v_t^i$$

While c_t^i and x_t^i are the observed consumption and shock outcomes, let us call c^{*i} and x^{*i} the underlying distribution of consumption and shock outcomes. Assuming that we can estimate the underlying distribution of shocks, we are left needing to estimate the underlying distribution of consumption outcomes. In the previous section we used the observed average of log consumption, $\overline{\ln c^i}$, to calculate our measure of poverty, but we would really like to use the expected average of log consumption, $\overline{\ln c^{*i}} = \overline{\ln c^i} - (\overline{x^i} - \overline{x^{*i}})\hat{\beta}$.[2]

In this way our measure of vulnerability could be more forward looking. The estimates in the previous section would find the household which experienced victimization in a year when survey data was collected as more vulnerable than the household which, by chance, did not experience victimization in a survey year. By modeling the evolution of shocks we can predict which shocks certain households are likely to face in the future, predict how they would effect consumption, and then estimate vulnerability from those prediction.

This path does not seem very promising for a few reasons. First of all, we know very little about why shocks occur and would have a difficult time modeling their occurrence. We might observe that in every period a quarter of households, who appear otherwise indistinguishable, experience robbery. We might then assume that every household has a 25% chance of being victimized. But, it may actually be the case that households who do not own guns get victimized, and that we do not have data on gun ownership. We will only make measures of poverty and risk worse if we construct a very inaccurate model of shocks.

Second, explained idiosyncratic risk in the South Africa data is only a little more than 16% of vulnerability. These figures are even lower for the

[2] I assume that there are enough households so that the error in η_t and thus the aggregate risk component is small.

Bulgaria and Vietnam data, perhaps because they contain less data on idiosyncratic shocks. (The figures are 1% in Bulgaria and 0.2% in Vietnam.) The lion's share of risk is unexplained risk. This may either be because consumption measures contain a lot of measurement error, or it may be because data sets do not contain good measures of many of the shocks which households face. It seems that attempting to explain the error term may be a better path to take.

3.2 Explaining the Error Term

The above discussion suggests another way of trying to improve predictions of the future rather than merely using experiences from the past. As 56% of vulnerability is due to unexplained risk, we may do a better job of capturing potential future vulnerability to risk by modeling the heteroskedasticity in the error term.

Chaudhuri (2001) has made a serious effort to do just that in his estimates of vulnerability. He allows the unexplained portion of the variance of consumption to depend on observable fixed characteristics of the sample households. He suggests vulnerability could be estimated as follows

$$\hat{V} = U(z) - \int U(\hat{c}^i e^{\hat{\sigma}_\nu(z^i)\epsilon}) d\Phi(\epsilon)$$

where $\hat{c}_i$ is the estimate of the explained portion of consumption as estimated from equation (3), z_i are the fixed characteristics, and σ_ν is the standard deviation of the error term in equation (3). This can be estimated using FGLS.[3]

Rather than estimating each household's vulnerability based only on the observed outcomes for that household, one can use estimates of the standard deviation of the error for each household to estimate the distribution of all possible outcomes which the household could face and calculate a more forward-looking measure of vulnerability in this manner. This is relatively simple, with the main disadvantage being that it is no longer possible to separate out a piece which may be due to measurement error. In Section 4 we will show how this techniques can be used and elaborated on to model both the explained and the unexplained portions of consumption.

[3]Ligon (2004) further allows the disturbance term to be correlated across households.

3.3 Assuming Difference Stationarity

Up to now we have been assuming that consumption follows a stationary process. A weaker stationarity assumption than the one made thus far would be to assume that changes in consumption are stationary and mean zero over time. In this case consumption follows a random walk. To incorporate this we would change the estimating equation to one such as the following

$$\ln c_t^i = \ln c_{t-1}^i + \eta_t + x_t^i \beta + v_t^i. \tag{4}$$

Using this one could go on to model the evolution of the shock variables x as well as the error term v as discussed in the previous two subsections.

With three periods of data, as we have in the South Africa KIDS data, one could conduct a test for stationarity. Ligon & Schechter (2004) suggest that if the data exhibits non-stationarity and has very low levels of measurement error then best practice is to estimate vulnerability using equation (4). But, with even moderate levels of measurement error it is found that using equation (3) and subtracting the portion of vulnerability due to unexplained risk yields more accurate results.

Even when assuming that the unexplained risk is not due to measurement error, assuming difference stationarity is problematic when trying to design a forward-looking measure. One must do a much better job at predicting the evolution of aggregate shocks (η_t) and unobserved shocks (v_t^i), since these shocks not only affect consumption today (as they did in the previous section) but also consumption forever after.

3.4 Abandoning Stationarity

Stationarity and difference stationarity are both strong assumptions. We may think that we live in a world with deterministic growth or a world with asset-based poverty traps. In such cases current observed consumption will not be a good indicator of household welfare. The previous three suggestions for estimating vulnerability discussed above will lead to serious misestimation and mistargeting. Although estimating a dynamic measure of vulnerability would improve targeting greatly, it requires the specification and estimation of the dynamic model.

Elbers & Gunning (2006) focus on just this issue and construct a dynamic model in which the capital stock for each household, and thus the consumption level, has some steady state level which depends on the household's level of productivity (for example soil quality and education). They estimate the following dynamic version of the vulnerability measure

$$V = \mathrm{E}_0 \sum_{t=s}^{T+s} \beta^{t-s}[U(z) - U(c_t^i)].$$

Thus, as Elbers & Gunning (2006) note, those households which are currently vulnerable, perhaps due to bad luck, are not those who are structurally vulnerable in the long run.

Of course their estimate of vulnerability is highly dependent on the assumptions they make regarding the underlying dynamic environment. They assume a simple Ramsey growth model and then identify chronically vulnerable households. But if, for example, the underlying dynamics involve asset-based poverty traps, as described by Carter & Barrett (2006), than one would be misestimating vulnerability yet again.[4] For a truly forward looking measure of vulnerability one needs to first have a model which can predict the future. Given the huge difficulties involved in correctly modeling the consumption generating process using economic theory, as well as the danger for mistargeting if one models the process incorrectly, it seems that econometric techniques, rather than those involving economic theory, may be more appropriate in this situation.

4 Focus on Explaining the Error

In this section of the paper we suggest a method of combining the techniques used by Ligon & Schechter (2003) and Chaudhuri (2001) and we examine the differing results given by the different methods. First we test whether there is any aggregate risk which can be explained by group membership. Ligon and Schechter estimate equation (3). This assumes that all households face the same aggregate risk. On the other hand, we might expect that there is some shock which only urban households face, or which only uneducated households face.

One roundabout way for testing this is testing whether the variance of the errors for those different groups of households is different. A Goldfeld-Quandt test can be used for this. In this test, equation (3) is estimated once

[4] Antman & McKenzie (2006) give evidence that urban Mexican households do not face income-based poverty traps. They mention that rural asset-based poverty traps may be more likely.

for the urban households and once for the rural households. The test statistic is then $F[n_r - K, n_u - K] = \frac{e_r' e_r/(n_r-K)}{e_u' e_u/(n_u-K)}$. If the variances of the two groups are significantly different we can estimate the following equation

$$\ln c_{tg}^i = \alpha^i + \eta_{tg} + x_{tg}^i \beta + v_{tg}^i. \tag{5}$$

The main difference is that the time fixed effect is now allowed to vary over groups. This may soak up some, but perhaps not all, of the difference in variance between groups. [*Perhaps a less roundabout way of getting at this would be to test whether the means of the errors differ across groups using a t-test rather than testing whether the variances of the errors differ across groups using the Goldfeld-Quandt test.*]

[*We are in the process of carrying out this sort of analysis using the KIDS data dividing households into categories based on rural/urban, Indian/African, high assets/low assets, and high education/low education. Using these techniques we hope to be able to soak up much more of the unexplained risk.*]

Incorporating a time-group fixed effect, based on the groups which the data tells you are important for risk, increases aggregate risk and decreases unexplained risk. This is a good way to start decreasing the unexplained portion of risk in the Ligon-Schechter measure. Still, we may think that the heteroskedasticity in the error has a more general format. It is sensible to assume that the disturbance variance varies with a set of regressors such as urban, Indian, assets, assets squared, income, income squared, and so on.

We incorporate this by assuming that $v_t^{i2} = z_t^i \kappa + u_t^i$ where z are the variables (including a constant) which influence the heteroskedasticity. We use a Breusch-Pagan test to test for heteroskedasticity of this form. (This test is a Lagrange Multiplier test with the form $LM = \frac{1}{2}[g'Z(Z'Z)^{-1}Z'g]$ where $g_t^i = \frac{v_t^{i2}}{(v'v/n)^{-1}}$, g is the vector of g_t^i, and Z is the matrix of z_t^i. If this test finds heteroskedasticity of this form then we will reestimate the model incorporating this heteroskedasticity.

If we do find such heteroskedasticity, then our OLS results are inefficient. Instead, we can use two-step feasible generalized least squares (FGLS) to measure more efficient estimates of the coefficients in either equation (3) or (5). One might expect that the estimates of vulnerability using FGLS would further decrease the size of unexplained risk, although it is not obvious that this is the case.

[*We are also currently in the process of carrying out this estimation. We are calculating FGLS versions of both equations (3) or (5) and will compare* *estimates of vulnerability using both techniques with those from the original more common estimation technique.*]

One should note that none of the estimations we are suggesting thusfar will actually change the total value of vulnerability, they will only change the breakdown between unexplained and explained vulnerability. This is because in equation (1) ($V = U(z) - \mathrm{E}U(c^i)$) our estimate of $\mathrm{E}U(c^i)$ is based on observed outcomes for c^i. In order to get our more forward-looking measure of chronic vulnerability, we would like to predict future outcomes of c^i, not just use observed outcomes for c^i.

The last step we will take is that of trying to predict the future based on the predicted variance of consumption more in line with the techniques used by Chaudhuri (2001). Now, rather than using observed consumption outcomes to estimate $\mathrm{E}U(c^i)$, we will instead estimate expected consumption as $\mathrm{E}[\ln c^i_t | i, t, X^i_t] = \alpha^i + \eta_t + x^i_t\beta$ and the variance of log consumption as $\hat{V}[\ln c^i_t | i, t, X^i_t] = z^i_t\kappa$. [*We might want to do a multiplicative correction, as in Chaudhuri, Jalan, and Suryahadi so that the mean of predicted consumption equals the mean of actual consumption.*]

In the end, we will be comparing the results from five estimation strategies, each trying to take one step further to predicting future consumption patterns. First, the results using the original Ligon-Schechter (LS) techniques which use OLS. Second, the results using the LS technique but implementing using FGLS. Third and fourth, the results using the LS technique but adding in time-group fixed effects with OLS and FGLS respectively. Lastly, FGLS estimation with time-group fixed effects and prediction of future consumption outcomes based on the variance of log consumption.

5 Implications for Practice

On the one hand, to have a truly forward-looking measure of chronic vulnerability, one would need to construct an accurate economic model of the consumption generating process. On the other hand, one may not feel confident that one's vision of the future is the accurate one on which to base targeting. In such a situation it might make sense to estimate a dynamic measure of vulnerability using multiple models and then compare the populations identified as vulnerable using each model. On the other hand, estimating multiple dynamic models is quite demanding.

We suggest a different route for estimating more forward-looking measures of chronic vulnerability. In the South African data used here, as well as in other data sets, a large portion of static vulnerability is unexplained risk. In this paper we attempt to estimate more seriously the distribution of this unexplained risk. We show a static estimate of vulnerability, as well as four measures of vulnerability which incorporate this unexplained risk using different estimation techniques. We discuss how and why they differ. Given our findings, we suggest that [*one technique*] is a good way of estimating chronic vulnerability.

References

Alderman, H., Babita, M., Lanjouw, J. O., Lanjouw, P., Makhathat, N., Mohamed, A., Özler, B. & Qaba, O. (2000), Combining census and survey data to construct a poverty map of South Africa, *in* 'Measuring Poverty in South Africa', Pretoria: Statistics South Africa.

Antman, F. & McKenzie, D. (2006), 'Poverty traps and nonlinear income dynamics with measurement error and individual heterogeneity', *Journal of Development Studies* . Forthcoming.

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

Carter, M. R. & Maluccio, J. A. (2003), 'Social capital and coping with economic shocks: An analysis of stunting of South African children', *World Development* **31**(7), 1147–1163.

Chaudhuri, S. (2001), Empirical methods for assessing household vulnerability to poverty. Unpublished Manuscript.

Elbers, C. & Gunning, J. W. (2006), Poverty, risk, and accumulation: Propoor policies when dynamics matter. Unpublished Manuscript.

Ligon, E. (2004), Targeting and informal insurance, *in* S. Dercon, ed., 'Insurance Against Poverty', Oxford University Press.

Ligon, E. & Schechter, L. (2003), 'Measuring vulnerability', *The Economic Journal* **113**(486), 95–102.

Ligon, E. & Schechter, L. (2004), Evaluating different approaches to estimating vulnerability, Social Protection Discussion Paper 0410, World Bank.

May, J., Agüero, J., Carter, M. R. & Timæus, I. (2006), The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd wave: Methods, first findings and an agenda for future research. Unpublished Manuscript.

World Bank (2000), *Voices of the Poor*, World Development Report.