Insurance and Rural Welfare: What Can Panel Data Tell Us?

Elbers, Chris; Gunning, Jan Willem; Pan, Lei

Insurance and Rural Welfare: What Can Panel Data Tell Us?

Insurance and Rural Welfare:
What Can Panel Data Tell Us?

Chris Elbers[a], Jan Willem Gunning, and Lei Pan

Free University, Amsterdam and Tinbergen Institute

May, 2006

Running title: Insurance and rural welfare

Abstract: Assessing the scope for insurance in rural communities usually requires a structural model of household behavior under risk. One of the few empirical applications of such models is the study by Rosenzweig and Wolpin (1993) who conclude that Indian farmers in the ICRISAT villages would not benefit from the introduction of formal weather insurance. In this paper we investigate how models such as theirs can be estimated from panel data on production and assets. We show that if assets can take only a limited number of values the coefficients of the model cannot be estimated with reasonable precision. We also show that this can affect the conclusion that insurance would not be welfare improving.

[a] Corresponding author; Free University Amsterdam and Tinbergen Institute; address: FEWEB-VU, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; phone: +31205986143 ; fax: +31205986004 ; email: celbers@feweb.vu.nl

1 Introduction

Insecurity is a key aspect of poverty. There now is substantial evidence that exposure to risk (in the absence of well-functioning financial markets) is often reflected in very large volatility of a household's consumption over time (e.g. Baulch and Hoddinott, 2000; Dercon, 2005). There still is little clarity on the policy implications of such churning. For example, Rosenzweig and Wolpin (1993) argue that rural households need micro credit, but Dercon (2005) favors formal insurance arrangements. While many insurance schemes are now being piloted (including rainfall insurance in Ethiopia and option contracts on coffee prices in several African countries) very little is known of their costs and benefits relative to existing risk coping institutions. To assess new policy initiatives in this area we need to know more about their welfare effects.

In choosing risk-coping strategies households face, of course, a trade-off between the mean and the volatility of their income. For example, a household can reduce the volatility of its income through crop diversification, but it will thereby lower its mean income by foregoing the gains from specialization. The household thereby pays an implicit risk premium and this must be taken into account in the evaluation of policy interventions, such as micro credit, price stabilization or formal insurance. The evaluation should assess not only the change in volatility as a result of the intervention, but also the change in the (implicit) premium.

Ideally, the scope for insurance can be evaluated by comparing villages with and without insurance where treatment villages have been selected randomly and the

distribution of shocks is known. Alternatively, if locations differ in risk exposure this can be exploited as a natural experiment to infer how behavior would change under insurance. Clearly, the villages must then be similar in all other respects. Matching or differencing offers some scope for relaxing this requirement but only if all relevant variables are observed. In practice unobserved heterogeneity is likely to be a major problem. In that case there is no alternative to estimating a structural model, using panel data. Such a model can be used to derive how agents would behave if insurance were available, even if no agent had been observed in that situation. Obviously, this requires estimation of the model's structural parameters. An important advantage of this procedure is that the researcher does not need to know the distribution of the shocks to which the households are exposed; estimates of the parameters characterizing that distribution will be generated as part of the estimation procedure.

Lucas (2003) suggested on the basis of a back-of-envelope calculation that insurance could not have a substantial effect on growth. However, he considered a situation with much less risk than is common in many developing countries. Unfortunately, for developing countries there are few empirical studies.

A notable exception is the famous paper by Rosenzweig and Wolpin (1993) (henceforth RW). They estimated a structural model using the ICRISAT data collected in three Indian villages. The model describes investment behavior under risk where investment in bullocks is the key decision variable. Bullock ownership can take three values: 0, 1 or 2. The model is estimated from data on production and bullock ownership, but not on consumption. RW stressed that the villagers in

their sample participated in an informal insurance arrangement which established a floor under their consumption level. They found that the introduction of actuarially fair insurance (AFI) would not be welfare improving: households were already sufficiently protected through informal insurance. (Strictly speaking RW could draw no such conclusion. In their analysis a household that would switch from the informal arrangement to AFI would still have to pay the implicit premium of the informal arrangement. Clearly, this biases the comparison in favor of the *status quo*.) Rosenzweig and Binswanger (1993) using the same data reached the opposite conclusion: the poorer ICRISAT farmers would benefit greatly from the introduction of insurance. Average profits in the bottom wealth quartile would increase by about one third for a reduction in weather risk of one standard deviation. However, their conclusion was not based on a dynamic model: a household's total wealth was taken as given. Elbers, Gunning and Kinsey (2005) estimated a structural model for smallholder households in Zimbabwe. They found a massive effect of the introduction of AFI: on average in their sample households would accumulate a capital stock (cattle) twice as large (over a 50-year period) as in the absence of insurance (when consumption smoothing is the only risk coping strategy available to them). As in RW this model is estimated on the basis of production and capital stock data but in this study the capital stock is a continuous variable.

The use of a discrete concept of the capital stock is intuitively appealing: cattle ownership is likely to be less subject to measurement error than, say, consumption data; the number of cows or bullocks is a discrete variable which is easily

remembered. However, this advantage may well come at very high econometric costs, as we show in this paper. The question we investigate is whether the RW research agenda is feasible in the sense that robust conclusions about the welfare effects of insurance can be derived from a structural model estimated from production and capital stock data if the capital stock can take only a limited number of values. Our approach is to specify a simplified version of the RW model; to use the model as data generating mechanism; to find out how accurately we can recover the underlying behavioral parameters by estimation, given the values of all other parameters and given the true model specification; and, finally, to use the findings to assess the robustness of the policy conclusion on the desirability of the introduction of insurance.

The structure of the paper is as follows. In the next section we show under what conditions the use of limited dependent variables in this class of models may lead to large standard errors. In section 3 we specify the model and use simulation experiments to derive the distribution of the structural coefficients. We find, as expected, very large errors. It turns out that in this class of models the RW conclusion as to the welfare effects of insurance is not robust. Section 4 concludes.

2 Limited Dependent Variables

Consider the following deterministic intertemporal optimization problem

$$\max_{c_t, k_t} \sum_{t=o}^{\infty} \beta^t u(c_t)$$

subject to the constraints

$$c_t + k_{t+1} = f(k_t)$$
$$c_t, k_t \geq 0$$

k_0 given,

where c is consumption, k the capital stock, $u(c)$ the instantaneous utility function and β a discount factor. The usual interpretation of the model is that output $f(k_t)$ can be consumed or used as input into next period's production. Note the absence of a depreciation term $(1-\delta)k$: to get a more compact notation the production function $f(k)$ represents the real value of output plus the after-production value of the asset. A solution $\hat{k}_t$, $t = 1,...$ to the optimization problem can be characterized by an investment function φ, with

$$\hat{k}_{t+1} = \varphi(f(\hat{k}_t))$$
$$\hat{k}_0 = k_0$$

We will assume that the solution is unique. Note that the optimization problem is stationary, implying that φ does not depend on t.

There are several ways to extend the decision problem to a stochastic framework. One possibility is to allow for random shocks in $f(k)$ and to maximize the expected

value of summed discounted utility with respect to investment policies. If the shocks are serially independent the optimal policy is again an investment function of the form $\varphi(f(k))$.

In the deterministic case the functions $f(\cdot)$ and $\varphi(\cdot)$ can be identified from a sufficient number of different k_0 observations and subsequent observations on the capital stock and output. (Note that there is no need to observe consumption since it follows from $c_t = k_{t+1} + f(k_t)$.) The question arises whether observation of $f(k_t)$ (subject to shocks) and $\hat{k}_t$ is sufficient to recover the behavioral parameters of the process, namely the discount factor β and the parameters of $u(c)$. The answer is affirmative: by integrating the Euler conditions for an optimal accumulation path we can normally recover β as well as the utility function (over the relevant part of its domain and up to an affine transformation).[1] Knowledge of the behavioral parameters allows us to study counterfactual situations, such as a change in production function $f(k)$ or the introduction of insurance in a stochastic context.[2]

Now suppose that k_t is restricted to a limited number (n) of integer values.[3] Recovering the behavioral parameters of the model now becomes problematic. For example, in the deterministic case since there is a one to one mapping from k_t to $\hat{k}_t$ and $f(k_t)$ can only take n numbers, if the number of behavioral parameters exceeds n they cannot be recovered, irrespective of the number of observations

---

[1] Elbers and Gunning (2002).

[2] From a positive economics perspective this is the very reason for writing accumulation as an optimization problem instead of being satisfied with a purely descriptive function φ.

[3] This is similar to the situation studied by RW where asset 'k' is in fact a *vector* of stocks: bullocks, pump and calves of various ages, all integer valued. The number of bullocks can take only three values: 0, 1 and 2.

because no more than *n* values of *k* are observed. In the stochastic case (with $f(k)$ subject to shocks) output is a continuous variable while k is discrete. The investment function φ is now a step function characterized by threshold values $\overline{w}_i$ and corresponding asset levels a_i so that

$$w_0 = a_0 = 0$$

$$\hat{k}_{t+1} = a_i \text{, if } \overline{w}_i \leq f(k_t) < \overline{w}_{i+1}$$

Since the asset levels a_i are integer-valued they do not carry local information about the behavioral parameters; the parameters must therefore be determined from the threshold values $\overline{w}_i$. The number of threshold values therefore determines the number of parameters of the behavioral process that can be estimated.

As an example take the case where k settles on a steady state value of 2 and all households have initial values k_0 below this number. Then the data will at best allow the researcher to determine the thresholds $\overline{w}_1$ and $\overline{w}_2$.[4] Consequently only two parameters of the behavioral process can (normally) be estimated.

More information can be obtained if there is more heterogeneity between observed cases. For instance, if the time horizon of the optimization is finite (say, until the death of the agent) and agents differ in age, then the investment function φ and the threshold values $\overline{w}_i$ become age-dependent. Also, the production function $f(k)$ could differ across households, again leading to multiplication of observable threshold values. If such heterogeneity affects some (but not all) of the parameters it can be exploited in pooled estimation.

---

[4] It is possible that $a_1 = 2$ in which case only w_1 can estimated.

RW use both methods. They assume a finite decision horizon, leading to heterogeneity in asset holding policies across decision makers of different age. Irreversibility of the installation of a pump leads to a shift in the production function for pump owners[5].

Whether heterogeneity can indeed be used to solve the estimation problem is an open question.[6] RW in fact fix the value of one of the three behavioral parameters, the discount factor. Further, as we will show for a simplified version of their model, but with similar parameter values, heterogeneity introduced by pump ownership does little to improve the situation since conditionally on pump-less threshold values, behavioral parameters have almost no effect on pump-inclusive thresholds. Also, age-based heterogeneity has very little effect on investment behavior except for households with elderly heads; this gives such households inordinate weight in the estimation procedure.

We conclude that estimation of behavioral parameters underlying an accumulation process is inherently difficult if the asset involved can take (or is observed to take) only a few discrete values. The problem can be solved by putting additional constraints on the parameters, but any counterfactual analyses based on the estimated parameters must then be checked for robustness against such constraints. We also suggest that the identification problem can be solved if the asset becomes continuous.

---

[5] We do not measure the effect of irreversibility in this paper but only keep this assumption to be in line with RW.

[6] Rust (1994) provides detailed discussion of the identification of dynamic discrete choice models, presenting sufficient conditions under which the model can be identified and a simple and general estimation theory is available. One of the restrictions required in his discussion is that shocks enter into the utility function additively, which is violated in the RW model. As Rust (1994) mentions, no general estimation theory exists for this case.

3 Simulations

For our simulations we use a simplified version of the RW model. We first describe

the model and then discuss the simulations.

3.1 The model

In the simplified RW model each agent solves:

$$\max_{\varphi_\tau} \sum_{t=o}^{T_\tau} \beta^t E \frac{[\max(c_t - c_{\min}, 0)]^{1-\gamma}}{1-\gamma}$$

subject to:

$$c_t + p_k k_{t+1} + p_p (P_t - P_{t-1}) = w_t = \alpha(k_t) + \alpha_p P_t + p_k k_t + \varepsilon_t \qquad (1)$$

k_0 is given

$$k_{t+1} = \varphi_\tau(w_t) \qquad (2)$$

$$\varphi_\tau(w_t) = \begin{cases} 0, & w_t - p_b < c_{\min} \\ 0/1, & c_{\min} + p_b \leq w_t < c_{\min} + 2p_b \\ 0/1/2, & w_t > c_{\min} + 2p_b \end{cases} \qquad (3)$$

where the agent's instantaneous utility function is characterised by the parameter

$\gamma > 0$ ($\gamma \neq 1$), β is the discount factor, $c_{\min}$ is a minimum consumption level

supported by an informal insurance arrangement,[7] c is the level of consumption

---

[7] In this RW specification $c_{\min}$ has unfortunately a dual role: it is the level of consumption below
which a household cannot survive and also the level guaranteed by the informal insurance

before any support from the insurance arrangement, w is wealth at hand (available for consumption and investment), k is the capital stock (constrained to take the values 0, 1 or 2) with price p_k, P is a dummy variable indicating whether the household owns a pump (with price p_p) and ε is an income shock (a draw from a distribution known by the agent). Income shocks are independent over time. The function $\alpha(k)$ takes the values α_0, α_1, α_2 for $k = 0$, 1, 2 respectively; this picks up the effect of bullock ownership on income. The productivity of pump is captured by α_p. The age of the household head (which determines the time remaining until time T) is denoted by τ. In this formulation investment in pumps (which is irreversible) is exogenous (contrary to the RW formulation). Agents without a pump do not expect to acquire one. Expected utility is maximized with respect to the policy function φ_τ.

Note from (1) and (3) that the insurance arrangement will pay out only to households without cattle: a household with $c < c_{\min}$ will have to sell its cattle before it is entitled to consumption support since c is defined as the difference between wealth at hand (w) which includes the cattle asset and the investment.

The optimal policy function will depend on τ and P. Given these two parameters the function will involve two threshold values for wealth at hand, $\overline{w}_1$, $\overline{w}_2$ such that in

---

arrangement. It would be more natural to specify different parameters for these roles so that the welfare effect of a change in the protection offered by the insurance scheme can be analyzed.

each period the household chooses $k = 1$ for $\overline{w}_1 \leq w < \overline{w}_2$, $k = 2$ for $w \geq \overline{w}_2$ and $k = 0$ otherwise[8].

Like RW we assume that k_0, τ, P and the capital stock k_t are observed without error, but that instead of w_t the researcher observes $\widetilde{w}_t = w_t + \eta_t$, where η_t is measurement error, i.i.d. and independent of household shocks ε_t.

3.2 Simulations set-up

We impose the following values (based on the RW estimates) for the three structural coefficients: $c_{\min} = 1469$, $\beta = 0.95$, $\gamma = 0.964$. We set $p_k = 992$, $p_p = 6338$, $\alpha_0 = 0$, $\alpha_1 = 326$, $\alpha_2 = 1800$, $\alpha_p = 1795$ (based on RW estimates and the value used in their paper). The distribution of ε is normal with zero mean and $\sigma_\varepsilon = 2293$. Note that this is very high relative to the value of $c_{\min}$. The observation error η is also normally distributed with zero mean and standard deviation $\sigma_\eta = 427$. The maximum plan horizon T_0 is 70 periods. We solve the investment thresholds $\overline{w}$ by using backward recursion which is discussed in Rust (1994) as the main solution method for finite-horizon models. With these parameter values the investment thresholds $\overline{w}$ are virtually insensitive to age τ for all but the oldest households (see Figure 1). We have therefore eliminated age heterogeneity, imposing $T_0 = \infty$. For these parameter values we find thresholds $\overline{w}_1 = 2461$ and $\overline{w}_2 = 3453$ for

---

[8] There are two possible cases for the order of the thresholds. Besides the case used in this paper, the only possible case is the one mentioned in footnote 4. Since we do observe households with one bullock in the data, we exclude this case in this paper.

households without a pump and again $\overline{w}_1^P = 2461$ and $\overline{w}_2^P = 3453$ for households owning a pump.[9]

The set-up of the simulations is as follows. Each of thirty households is endowed with exogenous values k_0 and P_t. Next we generate a series of eight[10] shocks ε_t and measurement errors η_t for each household. The shocks are independent across households.[11] The measurement errors are applied to the true values w_t to generate the 'observed' values $\widetilde{w}_t$. The dataset now consists of a vector $\{k_1, \widetilde{w}_1, ..., k_s, \widetilde{w}_s \mid k_0, P_t\}$ for each household. The maximum likelihood estimates of the three behavioral parameters $\theta = (\gamma, c_{\min}, \beta)$ are derived for this data set (with all other parameters set at their true values). We then generate a new data set and a new set of estimates of θ. By repeating this procedure of data generation and estimation many times we generate the sampling distribution of the θ-estimators, given the values of all other parameters and the true underlying model specification.

Maximum Likelihood Estimation

Define $D_{1t} = 1$ if $k_t = 1$, and zero otherwise. Similarly for $D_{2t} = 1$ if $k_t = 2$. With this notation the likelihood contribution of a household is proportional to

$$\prod_{t=0}^{9} [\Pr(k_{t+1} = 0 \mid \tilde{w}_t, k_t, P_t)^{1-D_{1,t+1}-D_{2,t+1}} \times \Pr(k_{t+1} = 1 \mid \tilde{w}_t, k_t, P_t)^{D_{1,t+1}} \times \Pr(k_{t+1} = 2 \mid \tilde{w}_t, k_t, P_t)^{D_{2,t+1}}]$$

---

[9] Note that the threshold values are insensitive to pump ownership.

[10] Eight is also the number of periods in RW. RW do not mention the number of households, but the ICRISAT data set contains 30 medium-size households with two or more observations. RW use the data from this medium-size group of households.

[11] This is, of course, a simplification; in practice there will be covariance, e.g. in the case of weather shocks.

The three probabilities after the product sign are derived from a normal

distribution [12] $N((\varepsilon + \eta)\sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_\eta^2), \sigma_\varepsilon^2 \sigma_\eta^2 / \sqrt{\sigma_\varepsilon^2 + \sigma_\eta^2}$ with cumulative distribution

function Γ. For instance, if the household has no pump

$$\Pr(k_{t+1} = 0 \mid \widetilde{w}_t, k_t, P_t = 0) = \Gamma(\overline{w}_{1t} - \alpha_0) \qquad (4)$$

Note that the likelihood depends only indirectly, through the threshold values $\overline{w}$,

on the parameters θ. Maximizing the likelihood therefore involves computing the

threshold values as a function of θ. To reduce the computational burden we have

calculated the thresholds on a grid of parameters θ. Linear interpolation [13] of these

threshold values on the grid values of θ is then used to approximate the threshold

values for non-grid values of θ. The interpolated values turn out to be highly

accurate approximations to the exact threshold values. We then substitute the

interpolated threshold values $\hat{w}(\theta)$ in the likelihood function. For instance, in

equation (2) we get

$$\hat{\Pr}(k_{t+1} = 0 \mid \widetilde{w}_t, k_t, P_t = 0) = \Gamma(\hat{w}_{1t}(\theta) - \alpha_0)$$

It is now straightforward to maximize the likelihood with respect to θ. In all cases

we have used the true parameter values ($\beta = 0.95$, $\gamma = 0.964$, $c_{\min} = 1469$) as initial

point for the optimizing algorithm. Moreover, we have restricted parameters to the

bounds imposed of the grid of parameters for which the true thresholds have been

computed, i.e. intervals $(0.59, 0.99)$ for β, $(0.899, 0.989)$ for γ and $(0, 2997)$ for $c_{\min}$.

The bounds have been chosen to make the thresholds estimated precisely. Since

---

[12] Note that $\varepsilon + \eta = \widetilde{w}_t - \alpha(k_t) - \alpha_p P_t - p_k k_t$.

[13] The interpolation method used is a trivariate version of the "Four Point Formula" in Abramowitz and Stegun (1972, p. 882).

the estimates of β, γ and $c_{\min}$ should be close to their true values if they could be estimated from the model the choice of different bounds do not make difference to the estimation of the parameters.

3.3 Results

With the above choice of model and parameters it turns out that the threshold values $\overline{w}_{1,2}$ are virtually insensitive to parameters β and γ. Hence it is impossible to estimate these parameters with any accuracy. Given the typical size of household shocks and only 30 households, the influence of β and γ on the likelihood function is almost absent. We find that γ tends to settle on one of the bounds and β tends to stay very close to its initial value. Figure 2 is a typical scatter plot of the joint (β, γ) sampling distribution, based on 100 simulations.

On the other hand, $c_{\min}$ can be estimated fairly accurately. Figure 3 displays the sampling distribution of the estimator for $c_{\min}$. The mean and standard deviation determined from 100 simulations are 1444.7 and 80.3.

Recall that the heterogeneity in pump ownership raises the number of thresholds above the number of parameters so that the problem discussed in section 2 of having to recover three coefficients from two threshold values does not arise. However, it turns out that the extra pair of thresholds do not convey sufficient additional information: as may be seen from Figures 4 and 5 the thresholds for pump-owning households are almost insensitive to the behavioral parameters given the corresponding thresholds for the other households. Hence the second

pair of thresholds convey little extra information on the parameters. In fact, we find that the thresholds are almost exclusively determined by the value of $c_{\min}$, so that the combined figures trace essentially a one-dimensional sub-set in the space of threshold values. This is why only a single parameter can be estimated with reasonable precision.

What are the implications for the robustness of the conclusion that actuarially fair insurance would not raise welfare? We investigate this by calculating ΔW, the net increase in welfare (i.e. the expected value of discounted utility) as a result of replacing the informal arrangement supporting $c_{\min}$ with AFI.[14] We do this for each of six household types, defined by whether they have (initially) a pump and whether they start with 0, 1 or 2 bullocks. Table 1 shows ΔW for each of these household types and for various values of β in the $(0.59, 0.99)$ range, and putting all other parameters to their true value.[15] The values shown in the Table are calculated under the assumption that no household will acquire a pump: the first three household types remain pumpless throughout.

The Table shows that the first two types of households (those who initially have no pump and at most one bullock) would not be willing to give up the informal insurance arrangement (which guarantees a consumption level $c_{\min}$) in exchange

---

[14] Recall that this measure is biased since (as RW recognise) the cost of the informal arrangement would be reflected in a premium which would not have to be paid under AFI. Since RW do not know this cost (which is reflected in $\alpha(k)$) they ignore it in the comparison: hence under AFI the household continues to pay the same premium. This is a major issue: given the choice between AFI (i.e. constant consumption) and an arrangement where the household would receive positive shocks but (as a result of the $c_{\min}$ floor) no negative shocks it might well prefer the latter if the two schemes did not differ in cost. This would, obviously, say nothing about the desirability of AFI. Rather, it would reflect the failure to model explicitly the cost of the $c_{\min}$ scheme.

[15] The welfare changes can only be compared within rows since different rows correspond to different preferences.

for AFI. The reason is simple: under AFI they would no longer enjoy the positive income shocks which they experience under the (asymmetric) informal insurance arrangement. For richer households this advantage of the c_{min} arrangement is offset by a larger difference between c_{min} and mean consumption: the informal arrangement gives them little downward protection. As a result, households with pumps or with two bullocks would switch to AFI. Clearly, the aggregate effect depends on the distribution of the population over household types and on the discount factor. In the Table we show a particular distribution. Under this distribution the change in aggregate welfare depends on the value of the discount factor. For low values of β (i.e. a high discount rate) the net effect is negative, for values of 0.79 or higher it is positive: AFI would be accepted.

This example underestimates the case for AFI, for two reasons. First, as noted above, under AFI the implicit premium of the informal insurance would no longer have to be paid but this is not taken into account. Secondly, we have treated pump investment as exogenous. In the RW world pump investment is endogenous. A pumpless household might receive (at some future date) a positive shock large enough to enable it to buy a pump. At that stage insurance would become attractive. Therefore, if households were not forced to adopt insurance now or stay with the informal arrangement forever, but were instead offered the *option* of switching to insurance welfare then gains might be positive even for the first two types of households.

Since β cannot be estimated with precision, Table 1 implies that any policy conclusion on the welfare effect of introducing formal insurance will not be robust.

If β is estimated (in a relatively small sample) it may easily settle on one of the two boundary values (with opposite policy conclusions). Conversely, if β is fixed (as in RW, who set $\beta = 0.95$) the policy conclusion is thereby fixed as well.

4 Conclusion

There is a renewed interest in insurance mechanisms to assist rural households in risk coping. It is rarely possible to evaluate such interventions through (quasi) experimental evaluation methods. In principle estimating a structural model of household behavior under risk (using panel data) is a viable alternative. Rosenzweig and Wolpin (1993) applied this approach to the ICRISAT data and found that insurance would not be welfare improving. In this paper we have investigated the robustness of this policy conclusion. We have considered a class of models of household behavior under risk where assets can take only a small number of values. This severely restricts the scope for estimating structural coefficients. In the RW case heterogeneity could solve this problem but we have shown (for a simplified version of their model) that in small samples two of the coefficients cannot be estimated with reasonable precision in spite of this heterogeneity. Since the policy conclusion on the desirability of introducing formal insurance is sensitive to the value of these coefficients, the conclusion is probably not robust. This does not mean that we cannot use panel data to assess the scope for insurance. Rather, it implies that if asset data indeed take only a small number

of values (relative to the number of parameters to be estimated) then estimation requires heterogeneity which (unlike the heterogeneity allowed for by RW) leads to *independent* variation in threshold values. The procedure we have described can easily establish whether this condition is satisfied. Another way to solve the problem is by treating asset variable(s) continuous, *e.g.*, by us ing livestock (an aggregate of cattle, goats, sheep *etc.*) rather than bullocks as the asset variable.

References

Abramowitz, Milton and Irene A. Stegun (1972) (eds), *Handbook of Mathematical Functions*, Dover Publications, New York.

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

Dercon, Stefan (2005) Risk, Insurance and Poverty: a Review, in Stefan Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER.

Elbers, Chris and Jan Willem Gunning (2002) Growth Regressions and Economic Theory, Tinbergen Institute Discussion paper 034/2.

Elbers, Chris, Jan Willem Gunning and Bill Kinsey (2005) Growth and Risk: Methodology and Micro Evidence, Tinbergen Institute Discussion paper.

Lucas, Robert E., Jr. (2003) Macroeconomic Priorities, *American Economic Review*, vol. 93, pp. 1-14.

Rosenzweig, Mark R. and Hans Binswanger (1993) Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments, *Economic Journal*, vol. 103, pp. 56-78.

Rosenzweig, Mark R. and Kenneth Wolpin (1993) Credit Market Constraints,

Consumption Smoothing, and the Accumulation of Durable Production Assets in

Low-Income Countries: Investment in Bullocks in India, *Journal of Political*

Economy, vol.101, pp. 223-244.

Rust (1994) Structural estimation of Markov decision processes, in Handbook of

Econometrics, vol. 4, ed. by R. Engle and D. McFadden, Amsterdam: North

Holland.

Endowment							
initial k	0	1	2	0	1	2	
pump	0	0	0	1	1	1	
discount							weighted
factor							effect
.590	-25.8	-35.6	27.5	26.0	17.7	8.5	loss
.615	-27.8	-37.6	30.4	27.0	18.6	9.4	loss
.640	-30.0	-40.0	33.8	28.1	19.7	10.4	loss
.665	-32.5	-42.7	37.7	29.4	20.9	11.6	loss
.690	-35.5	-45.8	42.3	30.8	22.3	13.0	loss
.715	-39.0	-49.5	47.8	32.4	23.9	14.7	loss
.740	-43.2	-53.9	54.5	34.2	25.8	16.7	loss
.765	-48.4	-59.2	62.6	36.3	28.0	19.1	loss
.790	-54.7	-65.7	72.8	38.9	30.8	22.2	gain
.815	-62.8	-74.1	86.0	42.1	34.2	26.2	gain
.840	-73.5	-85.0	103.4	46.2	38.7	31.5	gain
.865	-88.2	-99.8	127.4	51.5	44.8	38.8	gain
.890	-109.6	-121.5	162.7	59.1	53.5	49.6	gain
.915	-143.5	-155.6	218.7	70.8	67.1	66.7	gain
.940	-203.7	-216.0	318.6	91.1	91.2	97.3	gain
.965	-328.8	-341.4	526.8	132.8	141.1	161.6	gain
.990	-640.5	-653.4	1047.0	236.6	266.2	323.6	gain
weights	1.8	1	1	1	1	1	

Table 1: Welfare Effect of Insurance

Figure 1: Investment thresholds $\overline{w}_1$ (below) and $\overline{w}_2$ (above) for households without a pump.

Figure 2: Sampling Distribution of Behavioral Parameters (β , γ)

Figure 3: Sampling Distribution of the c_{min} Estimator

Figure 4: Investment thresholds: $\bar{w}_1$ (vertically) against $\bar{w}_1^p$ (horizontally)

Figure 5: Investment thresholds: $\bar{w}_2$ (vertically) against $\bar{w}_2^p$ (horizontally)

25

Insurance and Rural Welfare:
What Can Panel Data Tell Us?

Chris Elbers[a], Jan Willem Gunning, and Lei Pan

Free University, Amsterdam and Tinbergen Institute

Revised 25 September 2006

Running title: Insurance and rural welfare

Abstract: Assessing the scope for insurance in rural communities usually requires a structural model of household behavior under risk. One of the few empirical applications of such models is the study by Rosenzweig and Wolpin (1993) who conclude that Indian farmers in the ICRISAT villages would not benefit from the introduction of formal weather insurance. In this paper we investigate 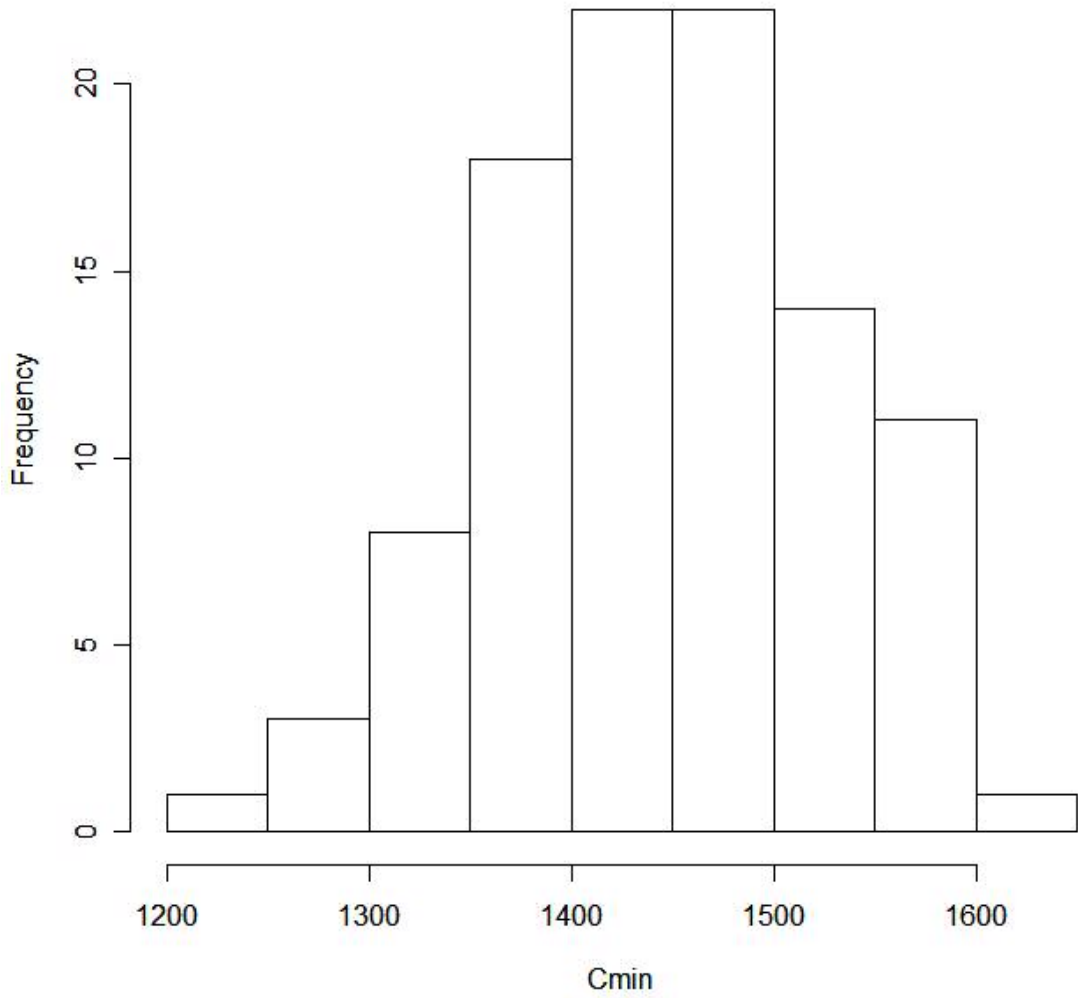how models such as theirs can be estimated from panel data on production and assets. We show that if assets can take only a limited number of values the coefficients of the model cannot be estimated with reasonable precision. We also show that this can affect the conclusion that insurance would not be welfare improving.

[a] Corresponding author; Free University Amsterdam and Tinbergen Institute; address: FEWEB-VU, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; phone: +31205986143 ; fax: +31205986004 ; email: celbers@feweb.vu.nl

1 Introduction

Insecurity is a key aspect of poverty. There now is substantial evidence that exposure to risk (in the absence of well-functioning financial markets) is often reflected in very large volatility of a household's consumption over time (e.g. Baulch and Hoddinott, 2000; Dercon, 2005). There still is little clarity on the policy implications of such churning. For example, Rosenzweig and Wolpin (1993) argue that rural households need micro credit, but Dercon (2005) favors formal insurance arrangements. While many insurance schemes are now being piloted (including rainfall insurance in Ethiopia and option contracts on coffee prices in several African countries) very little is known of their costs and benefits relative to existing risk coping institutions. To assess new policy initiatives in this area we need to know more about their welfare effects.

In choosing risk-coping strategies households face, of course, a trade-off between the mean and the volatility of their income. For example, a household can reduce the volatility of its income through crop diversification, but it will thereby lower its mean income by foregoing the gains from specialization. The household thereby pays an implicit risk premium and this must be taken into account in the evaluation of policy interventions, such as micro credit, price stabilization or formal insurance. The evaluation should assess not only the change in volatility as a result of the intervention, but also the change in the (implicit) premium. Households can also use consumption smoothing by accumulating or decumulating assets to cope with

1

risk.[1] Here the implicit premium is the loss of production as a result of using assets for this purpose.

Ideally, the scope for insurance can be evaluated by comparing villages with and without insurance where treatment villages have been selected randomly and the distribution of shocks is known. Alternatively, if locations differ in risk exposure this can be exploited as a natural experiment to infer how behavior would change under insurance. Clearly, the villages must then be similar in all other respects. Matching or differencing offers some scope for relaxing this requirement but only if all relevant variables are observed. In practice unobserved heterogeneity is likely to be a major problem. In that case there is no alternative to estimating a structural model, using panel data. Such a model can be used to derive how agents would behave if insurance were available, even if no agent had been observed in that situation. Obviously, this requires estimation of the model's structural parameters. An important advantage of this procedure is that the researcher does not need to know the distribution of the shocks to which the households are exposed; estimates of the parameters characterizing that distribution will be generated as part of the estimation procedure.

Lucas (2003) suggested on the basis of a back-of-envelope calculation that insurance could not have a substantial effect on growth. However, he considered a situation with much less risk than is common in many developing countries. Unfortunately, for developing countries there are few empirical studies.

A notable exception is the famous paper by Rosenzweig and Wolpin (1993) (henceforth RW). They estimated a structural model using the ICRISAT data

---

[1] See e.g. Tanner (1997) for US evidence and Elbers *et al.* (forthcoming) for Zimbabwe evidence.

collected in three Indian villages. The model describes investment behavior under risk where investment in bullocks is the key decision variable. Bullock ownership can take three values: 0, 1 or 2. The model is estimated from data on production and bullock ownership, but not on consumption. RW stressed that the villagers in their sample participated in an informal insurance arrangement which established a floor under their consumption level. They found that the introduction of actuarially fair insurance (AFI) would not be welfare improving: households were already sufficiently protected through informal insurance. Rosenzweig and Binswanger (1993) using the same data reached the opposite conclusion: the poorer ICRISAT farmers would benefit greatly from the introduction of insurance. Average profits in the bottom wealth quartile would increase by about one third for a reduction in weather risk of one standard deviation. However, their conclusion was not based on a dynamic model: a household's total wealth was taken as given. Elbers, Gunning and Kinsey (forthcoming) estimate a structural model for smallholder households in Zimbabwe. They found a massive effect of the introduction of AFI: on average in their sample households would accumulate a capital stock (cattle) twice as large (over a 50-year period) as in the absence of insurance (when consumption smoothing is the only risk coping strategy available to them). As in RW this model is estimated on the basis of production and capital stock data but in this study the capital stock is a continuous variable.

In many rural economies the key asset of a household is livestock, e.g. one or two bullocks. The use of a discrete concept of the capital stock is therefore a natural choice. In addition, it is well-known that households recall their cattle ownership

quite accurately. An implication is that observing the number of cattle is likely to be less subject to measurement error than, say, consumption. However, this advantage may well come at very high econometric costs, as we show in this paper. The question we investigate is whether the RW research agenda is feasible in the sense that robust conclusions about the welfare effects of insurance can be derived from a structural model estimated from production and capital stock data if the capital stock can take only a limited number of values. Our approach is to specify a simplified version of the RW model; to use the model as data generating mechanism; to find out how accurately we can recover the underlying behavioral parameters by estimation, given the values of all other parameters and given the true model specification; and, finally, to use the findings to assess the robustness of the policy conclusion on the desirability of the introduction of insurance.

The structure of the paper is as follows. In the next section we show under what conditions the use of limited dependent variables in this class of models may lead to large standard errors. In section 3 we specify the model and use simulation experiments to derive the distribution of the structural coefficients. We find, as expected, very large errors. It turns out that in this class of models the RW conclusion as to the welfare effects of insurance is not robust. Section 4 concludes.

2 Limited Dependent Variables

Consider the following deterministic intertemporal optimization problem

$$\max_{c_t, k_t} \sum_{t=o}^{\infty} \beta^t u(c_t)$$

subject to the constraints

$$c_t + k_{t+1} = f(k_t)$$
$$c_t, k_t \geq 0$$

k_0 given,

where c is consumption, k the capital stock, $u(c)$ the instantaneous utility function and β a discount factor. The usual interpretation of the model is that output $f(k_t)$ can be consumed or used as input into next period's production. Note the absence of a depreciation term $(1-\delta)k$: to get a more compact notation the production function $f(k)$ represents the real value of output plus the after-production value of the asset. A solution $\hat{k}_t$, $t = 1,...$ to the optimization problem can be characterized by an investment function φ, with

$$\hat{k}_{t+1} = \varphi(f(\hat{k}_t))$$
$$\hat{k}_0 = k_0$$

We will assume that the solution is unique. Note that the optimization problem is stationary, implying that φ does not depend on t.

There are several ways to extend the decision problem to a stochastic framework. One possibility is to allow for random shocks in $f(k)$ and to maximize the expected

5

value of summed discounted utility with respect to investment policies. If the shocks are serially independent the optimal policy is again an investment function of the form $\varphi(f(k))$.

In the deterministic case the functions $f(\cdot)$ and $\varphi(\cdot)$ can be identified from a sufficient number of different k_0 observations and subsequent observations on the capital stock and output. (Note that there is no need to observe consumption since it follows from $c_t = k_{t+1} + f(k_t)$.) The question arises whether observation of $f(k_t)$ (subject to shocks) and $\hat{k}_t$ is sufficient to recover the behavioral parameters of the process, namely the discount factor β and the parameters of $u(c)$. The answer is affirmative: by integrating the Euler conditions for an optimal accumulation path we can normally recover β as well as the utility function (over the relevant part of its domain and up to an affine transformation). [2] Knowledge of the behavioral parameters allows us to study counterfactual situations, such as the introduction of insurance.[3]

Now suppose that k_t is restricted to a limited number (n) of integer values.[4] Recovering the behavioral parameters of the model now becomes problematic. For example, in the deterministic case since there is a one to one mapping from k_t to $\hat{k}_t$ and $f(k_t)$ can only take n numbers, if the number of behavioral parameters exceeds n they cannot be recovered, irrespective of the number of observations

---

[2] Elbers and Gunning (2002).

[3] From a positive economics perspective this is the very reason for writing accumulation as an optimization problem instead of being satisfied with a purely descriptive function φ.

[4] This is similar to the situation studied by RW where asset 'k' is in fact a *vector* of stocks: bullocks, pump and calves of various ages, all integer valued. The number of bullocks can take only three values: 0, 1 and 2.

because no more than *n* values of *k* are observed. In the stochastic case (with $f(k)$ subject to shocks) output is a continuous variable while k is discrete. The investment function φ is now a step function characterized by threshold values $\overline{w}_i$ and corresponding asset levels a_i so that

$$w_0 = a_0 = 0$$

$$\hat{k}_{t+1} = a_i, \text{ if } \overline{w}_i \leq f(k_t) < \overline{w}_{i+1}$$

Since the asset levels a_i are integer-valued they do not carry local information about the behavioral parameters; the parameters must therefore be determined from the threshold values $\overline{w}_i$. The number of threshold values therefore determines the number of parameters of the behavioral process that can be estimated.

As an example take the case where k settles on a steady state value of 2 and all households have initial values k_0 below this number. Then the data will at best allow the researcher to determine the thresholds $\overline{w}_1$ and $\overline{w}_2$.[5] Consequently only two parameters of the behavioral process can (normally) be estimated.

More information can be obtained if there is more heterogeneity between observed cases. For instance, if the time horizon of the optimization is finite (say, until the death of the agent) and agents differ in age, then the investment function φ and the threshold values $\overline{w}_i$ become age-dependent. Also, the production function $f(k)$ could differ across households, again leading to multiplication of observable threshold values. If such heterogeneity affects some (but not all) of the parameters it can be exploited in pooled estimation.

---

[5] It is possible that $a_1 = 2$ in which case only w_1 can estimated.

RW use both methods. They assume a finite decision horizon, leading to heterogeneity in asset holding policies across decision makers of different age. Irreversibility of the installation of a pump leads to a shift in the production function for pump owners.

Whether heterogeneity can indeed be used to solve the estimation problem is an open question.[6] RW in fact fix the value of one of the three behavioral parameters, the discount factor. Further, as we will show for a simplified version of their model, but with similar parameter values, heterogeneity introduced by pump ownership does little to improve the situation since conditionally on pump-less threshold values, behavioral parameters have almost no effect on pump-inclusive thresholds. Also, age-based heterogeneity has very little effect on investment behavior except for households with elderly heads; this gives such households inordinate weight in the estimation procedure.

We conclude that estimation of behavioral parameters underlying an accumulation process is inherently difficult if the asset involved can take only a few discrete values. The problem can be solved by putting additional constraints on the parameters, but any counterfactual analyses based on the estimated parameters must then be checked for robustness against such constraints. We also suggest that the identification problem can be solved if the asset becomes continuous.

---

[6] Rust (1994) provides detailed discussion of the identification of dynamic discrete choice models, presenting sufficient conditions under which the model can be identified and a simple and general estimation theory is available. One of the restrictions required in his discussion is that shocks enter into the utility function additively, which is violated in the RW model. As Rust (1994) mentions, no general estimation theory exists for this case.

3 Simulations

For our simulations we use a simplified version of the RW model.[7] We first describe the model and then discuss the simulations.

3.1 The model

In the simplified RW model each agent solves:

$$\max_{\varphi_\tau} \sum_{t=o}^{T_\tau} \beta^t E \frac{[\max(c_t - c_{\min}, 0)]^{1-\gamma}}{1-\gamma}$$

subject to:

$$c_t + p_k k_{t+1} + p_p(P_t - P_{t-1}) = w_t = \alpha(k_t) + \alpha_p P_t + p_k k_t + \varepsilon_t \qquad (1)$$

k_0 is given

$$k_{t+1} = \varphi_\tau(w_t) \qquad (2)$$

$$\varphi_\tau(w_t) = \begin{cases} 0, & w_t - p_k < c_{\min} \\ 0/1, & c_{\min} + p_k \leq w_t < c_{\min} + 2p_k \\ 0/1/2, & c_{\min} + 2p_k \leq w_t \end{cases} \qquad (3)$$

where the agent's instantaneous utility function is characterised by the parameter $\gamma > 0$ ($\gamma \neq 1$), β is the discount factor, $c_{\min}$ is a minimum consumption level supported by an informal insurance arrangement,[8] c is the level of consumption

---

[7] In this model consumption smoothing is the only coping mechanism. Maitra (2001) finds evidence of another coping mechanism, changes in labour supply.

[8] In this RW specification $c_{\min}$ has unfortunately a dual role: it is the level of consumption below which a household cannot survive and also the level guaranteed by the informal insurance arrangement. It would be more natural to specify different parameters for these roles so that the welfare effect of a change in the protection offered by the insurance scheme can be analyzed.

before any support from the insurance arrangement, w is wealth at hand (available for consumption and investment), k is the capital stock (constrained to take the values 0, 1 or 2) with price p_k, P is a dummy variable indicating whether the household owns a pump (with price p_p) and ε is an income shock (a draw from a distribution known by the agent). Income shocks are independent over time. The function $\alpha(k)$ takes the values α_0, α_1, α_2 for $k = 0$, 1, 2 respectively; this picks up the effect of bullock ownership on income. The productivity of pumps is captured by α_p. The age of the household head (which determines the time remaining until time T) is denoted by τ. In this formulation investment in pumps (which is irreversible) is exogenous (contrary to the RW formulation). Agents without a pump do not expect to acquire one. Expected utility is maximized with respect to the policy function φ_τ.

Note from (1) and (3) that the insurance arrangement will pay out only to households without cattle: a household with $c < c_{min}$ will have to sell its cattle before it is entitled to consumption support since c is defined as the difference between wealth at hand (w) which includes the cattle asset and the investment.

The optimal policy function will depend on τ and P. Given these two parameters the function will involve two threshold values for wealth at hand, $\overline{w}_1$, $\overline{w}_2$ such that in each period the household chooses $k = 1$ for $\overline{w}_1 \le w < \overline{w}_2$, $k = 2$ for $w \ge \overline{w}_2$ and $k = 0$ otherwise[9].

---

[9] There are two possible cases for the order of the thresholds. Besides the case used in this paper, the only possible case is the one mentioned in footnote 4. Since we do observe households with one bullock in the data, we exclude this case in this paper.

Like RW we assume that k_0, τ, P and the capital stock k_t are observed without error, but that instead of w_t the researcher observes $\tilde{w}_t = w_t + \eta_t$, where η_t is measurement error, i.i.d. and independent of household shocks ε_t.

3.2 Simulations set-up

We impose the following values (based on the RW estimates) for the three structural coefficients: $c_{\min} = 1469$, $\beta = 0.95$, $\gamma = 0.964$. We set $p_k = 992$, $p_p = 6338$, $\alpha_0 = 0$, $\alpha_1 = 326$, $\alpha_2 = 1800$, $\alpha_p = 1795$ (based on RW estimates and the value used in their paper). The distribution of ε is normal with zero mean and $\sigma_\varepsilon = 2293$. Note that this is very high relative to the value of $c_{\min}$. The observation error η is also normally distributed with zero mean and standard deviation $\sigma_\eta = 427$. The maximum plan horizon T_0 is 70 periods. We solve the investment thresholds $\overline{w}$ by using backward recursion which is discussed in Rust (1994) as the main solution method for finite-horizon models. With these parameter values the investment thresholds $\overline{w}$ are virtually insensitive to age τ for all but the oldest households. We have therefore eliminated age heterogeneity, imposing $T_0 = \infty$. For these parameter values we find thresholds $\overline{w}_1 = 2461$ and $\overline{w}_2 = 3453$ for households without a pump and again $\overline{w}_1^P = 2461$ and $\overline{w}_2^P = 3453$ for households owning a pump.[10]

---

[10] Note that the threshold values are insensitive to pump ownership.

The set-up of the simulations is as follows. Each of thirty households is endowed with exogenous values k_0 and P_t. Next we generate a series of eight[11] shocks ε_t and measurement errors η_t for each household. The shocks are independent across households.[12] The measurement errors are applied to the true values w_t to generate the 'observed' values $\tilde{w}_t$. The dataset now consists of a vector $\{k_1, \tilde{w}_1, ..., k_s, \tilde{w}_s \mid k_0, P_t\}$ for each household. The maximum likelihood estimates of the three behavioral parameters $\theta = (\gamma, c_{\min}, \beta)$ are derived for this data set (with all other parameters set at their true values). We then generate a new data set and a new set of estimates of θ. By repeating this procedure of data generation and estimation many times we generate the sampling distribution of the θ-estimators, given the values of all other parameters and the true underlying model specification.

Maximum Likelihood Estimation

Define $D_{1t} = 1$ if $k_t = 1$, and zero otherwise. Similarly for $D_{2t} = 1$ if $k_t = 2$. With this notation the likelihood contribution of a household is proportional to

$$\prod_{t=0}^{9} [\Pr(k_{t+1} = 0 \mid \tilde{w}_t, k_t, P_t)^{1 - D_{1,t+1} - D_{2,t+1}} \times \Pr(k_{t+1} = 1 \mid \tilde{w}_t, k_t, P_t)^{D_{1,t+1}} \times \Pr(k_{t+1} = 2 \mid \tilde{w}_t, k_t, P_t)^{D_{2,t+1}}]$$

---

[11] Eight is also the number of periods in RW. RW do not mention the number of households, but the ICRISAT data set contains 30 medium-size households with two or more observations. RW use the data from this medium-size group of households.

[12] This is, of course, a simplification; in practice there will be covariance, e.g. in the case of weather shocks.

The three probabilities after the product sign are derived from a normal

distribution [13] $N((\varepsilon + \eta)\sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_\eta^2), \sigma_\varepsilon^2\sigma_\eta^2/\sqrt{\sigma_\varepsilon^2 + \sigma_\eta^2}$ with cumulative distribution

function Γ. For instance, if the household has no pump

$$\Pr(k_{t+1} = 0 \mid \tilde{w}_t, k_t, P_t = 0) = \Gamma(\overline{w}_{1t} - \alpha_0) \tag{4}$$

Note that the likelihood depends only indirectly, through the threshold values $\overline{w}$,

on the parameters θ. Maximizing the likelihood therefore involves computing the

threshold values as a function of θ. To reduce the computational burden we have

calculated the thresholds on a grid of parameters θ. Linear interpolation[14] of these

threshold values on the grid values of θ is then used to approximate the threshold

values for non-grid values of θ. The interpolated values turn out to be highly

accurate approximations to the exact threshold values. We then substitute the

interpolated threshold values $\hat{w}(\theta)$ in the likelihood function. For instance, in

equation (2) we get

$$\hat{\Pr}(k_{t+1} = 0 \mid \tilde{w}_t, k_t, P_t = 0) = \Gamma(\hat{w}_{1t}(\theta) - \alpha_0)$$

It is now straightforward to maximize the likelihood with respect to θ. In all cases

we have used the true parameter values ($\beta = 0.95$, $\gamma = 0.964$, $c_{\min} = 1469$) as initial

point for the optimizing algorithm. Moreover, we have restricted parameters to the

bounds imposed of the grid of parameters for which the true thresholds have been

computed, i.e. intervals $(0.59, 0.99)$ for β, $(0.899, 0.989)$ for γ and $(0, 2997)$ for $c_{\min}$.

The bounds have been chosen to make the thresholds estimated precisely. Since

---

[13] Note that $\varepsilon + \eta = \tilde{w}_t - \alpha(k_t) - \alpha_p P_t - p_k k_t$.

[14] The interpolation method used is a trivariate version of the "Four Point Formula" in Abramowitz and Stegun (1972, p. 882).

the estimates of β, γ and $c_{\min}$ should be close to their true values if they could be estimated from the model the choice of different bounds do not make difference to the estimation of the parameters.

3.3 Results

With the above choice of model and parameters it turns out that the threshold values $\overline{w}_{1,2}$ are virtually insensitive to parameters β and γ. Hence it is impossible to estimate these parameters with any accuracy. We find that γ tends to settle on one of the bounds and β tends to stay very close to its initial value. Figure 1 is a typical scatter plot of the joint (β, γ) sampling distribution, based on 100 simulations.

On the other hand, $c_{\min}$ can be estimated fairly accurately. The mean and standard deviation determined from 100 simulations are 1444.7 and 80.3.

Recall that the heterogeneity in pump ownership raises the number of thresholds above the number of parameters so that the problem discussed in section 2 of having to recover three coefficients from two threshold values does not arise. However, it turns out that the extra pair of thresholds do not convey sufficient additional information: as may be seen from Figures 2 and 3 the thresholds for pump-owning households are almost insensitive to the behavioral parameters given the corresponding thresholds for the other households. Hence the second pair of thresholds convey little extra information on the parameters. In fact, we find that the thresholds are almost exclusively determined by the value of $c_{\min}$, so that

the combined figures trace essentially a one-dimensional sub-set in the space of

threshold values. This is why only a single parameter can be estimated with

reasonable precision.

What are the implications for the robustness of the conclusion that actuarially fair

insurance would not raise welfare? We investigate this by calculating ΔW, the net

increase in welfare (i.e. the expected value of discounted utility) as a result of

replacing the informal arrangement supporting $c_{\min}$ with AFI.[15] We do this for each

of six household types, defined by whether they have (initially) a pump and

whether they start with 0, 1 or 2 bullocks. Table 1 shows ΔW for each of these

household types and for various values of β in the $(0.59, 0.99)$ range, and putting

all other parameters to their true value.[16] The values shown in the Table are

calculated under the assumption that no household will acquire a pump: the first

three household types remain pumpless throughout.

The Table shows that the first two types of households (those who initially have no

pump and at most one bullock) would not be willing to give up the informal

insurance arrangement (which guarantees a consumption level $c_{\min}$) in exchange

for AFI. The reason is simple: under AFI they would no longer enjoy the positive

income shocks which they experience under the (asymmetric) informal insurance

---

[15] Recall that this measure is biased since (as RW recognise) the cost of the informal arrangement would be reflected in a premium which would not have to be paid under AFI. Since RW do not know this cost (which is reflected in $\alpha(k)$) they ignore it in the comparison: hence under AFI the household continues to pay the same premium. This is a major issue: given the choice between AFI (i.e. constant consumption) and an arrangement where the household would receive positive shocks but (as a result of the $c_{\min}$ floor) no negative shocks it might well prefer the latter if the two schemes did not differ in cost. This would, obviously, say nothing about the desirability of AFI. Rather, it would reflect the failure to model explicitly the cost of the $c_{\min}$ scheme.

[16] The welfare changes can only be compared within rows since different rows correspond to different preferences.

arrangement. For richer households this advantage of the $c_{\min}$ arrangement is offset by a larger difference between $c_{\min}$ and mean consumption: the informal arrangement gives them little downward protection. As a result, households with pumps or with two bullocks would switch to AFI. Clearly, the aggregate effect depends on the distribution of the population over household types and on the discount factor. In the Table we show a particular distribution. Under this distribution the change in aggregate welfare depends on the value of the discount factor. For low values of β (i.e. a high discount rate) the net effect is negative, for values of 0.79 or higher it is positive: AFI would be accepted.

This example underestimates the case for AFI, for two reasons. First, as noted above, under AFI the implicit premium of the informal insurance would no longer have to be paid but this is not taken into account. Secondly, we have treated pump investment as exogenous. In the RW world pump investment is endogenous. A pumpless household might receive (at some future date) a positive shock large enough to enable it to buy a pump. At that stage insurance would become attractive. Therefore, if households were not forced to adopt insurance now or stay with the informal arrangement forever, but were instead offered the *option* of switching to insurance then welfare gains might be positive even for the first two types of households.

Since β cannot be estimated with precision, Table 1 implies that any policy conclusion on the welfare effect of introducing formal insurance will not be robust. If β is estimated (in a relatively small sample) it may easily settle on one of the two

boundary values (with opposite policy conclusions). Conversely, if β is fixed (as in

RW, who set $\beta = 0.95$) the policy conclusion is thereby fixed as well.

4 Conclusion

There is a renewed interest in insurance mechanisms to assist rural households in

risk coping. It is rarely possible to evaluate such interventions through (quasi)

experimental evaluation methods. In principle estimating a structural model of

household behavior under risk (using panel data) is a viable alternative.

Rosenzweig and Wolpin (1993) applied this approach to the ICRISAT data and

found that insurance would not be welfare improving. In this paper we have

investigated the robustness of this policy conclusion. We have considered a class

of models of household behavior under risk where assets can take only a small

number of values. This severely restricts the scope for estimating structural

coefficients. In the RW case heterogeneity could solve this problem but we have

shown (for a simplified version of their model) that in small samples two of the

coefficients cannot be estimated with reasonable precision in spite of this

heterogeneity. Since the policy conclusion on the desirability of introducing formal

insurance is sensitive to the value of these coefficients, the conclusion is probably

not robust. This does not mean that we cannot use panel data to assess the scope

for insurance. Rather, it implies that if asset data indeed take only a small number

of values (relative to the number of parameters to be estimated) then estimation

requires heterogeneity which (unlike the heterogeneity allowed for by RW) leads to *independent* variation in threshold values. The procedure we have described can easily establish whether this condition is satisfied. An alternative is to treat assets as continuous variables, *e.g* by using livestock (an aggregate of cattle, goats, sheep *etc.*) rather than bullocks as the capital stock.

18

References

Abramowitz, Milton and Irene A. Stegun (1972) (eds), *Handbook of Mathematical Functions*, Dover Publications, New York.

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

Dercon, Stefan (2005) Risk, Insurance and Poverty: a Review, in Stefan Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER.

Elbers, Chris and Jan Willem Gunning (2002) Growth Regressions and Economic Theory, Tinbergen Institute Discussion paper 034/2.

Elbers, Chris, Jan Willem Gunning and Bill Kinsey (forthcoming) Growth and Risk: Methodology and Micro Evidence, *World Bank Economic Review*.

Lucas, Robert E., Jr. (2003) Macroeconomic Priorities, *American Economic Review*, vol. 93, pp. 1-14.

Maitra, Pushkar (2001) Is Consumption Smooth at the Cost of Volatile Leisure? An Investigation of Rural India, *Applied Economics*, vol. 33, pp. 727-734.

Rosenzweig, Mark R. and Hans Binswanger (1993) Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments, *Economic Journal*, vol. 103, pp. 56-78.

Rosenzweig, Mark R. and Kenneth Wolpin (1993) Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India, *Journal of Political Economy*, vol.101, pp. 223-244.

Rust, John (1994) Structural estimation of Markov decision processes, in *Handbook of Econometrics*, vol. 4, ed. by R. Engle and D. McFadden, Amsterdam: North Holland.

Tanner, Evan (1997) Shifts in US Savings: Long-run Asset Accumulation versus Consumption Smoothing, *Applied Economics*, vol. 29, pp. 989-999.

Endowment							
initial k	0	1	2	0	1	2	
pump	0	0	0	1	1	1	
discount							weighted
factor							effect
.590	-25.8	-35.6	27.5	26.0	17.7	8.5	loss
.615	-27.8	-37.6	30.4	27.0	18.6	9.4	loss
.640	-30.0	-40.0	33.8	28.1	19.7	10.4	loss
.665	-32.5	-42.7	37.7	29.4	20.9	11.6	loss
.690	-35.5	-45.8	42.3	30.8	22.3	13.0	loss
.715	-39.0	-49.5	47.8	32.4	23.9	14.7	loss
.740	-43.2	-53.9	54.5	34.2	25.8	16.7	loss
.765	-48.4	-59.2	62.6	36.3	28.0	19.1	loss
.790	-54.7	-65.7	72.8	38.9	30.8	22.2	gain
.815	-62.8	-74.1	86.0	42.1	34.2	26.2	gain
.840	-73.5	-85.0	103.4	46.2	38.7	31.5	gain
.865	-88.2	-99.8	127.4	51.5	44.8	38.8	gain
.890	-109.6	-121.5	162.7	59.1	53.5	49.6	gain
.915	-143.5	-155.6	218.7	70.8	67.1	66.7	gain
.940	-203.7	-216.0	318.6	91.1	91.2	97.3	gain
.965	-328.8	-341.4	526.8	132.8	141.1	161.6	gain
.990	-640.5	-653.4	1047.0	236.6	266.2	323.6	gain
weights	1.8	1	1	1	1	1	

Table 1: Welfare Effect of Insurance

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
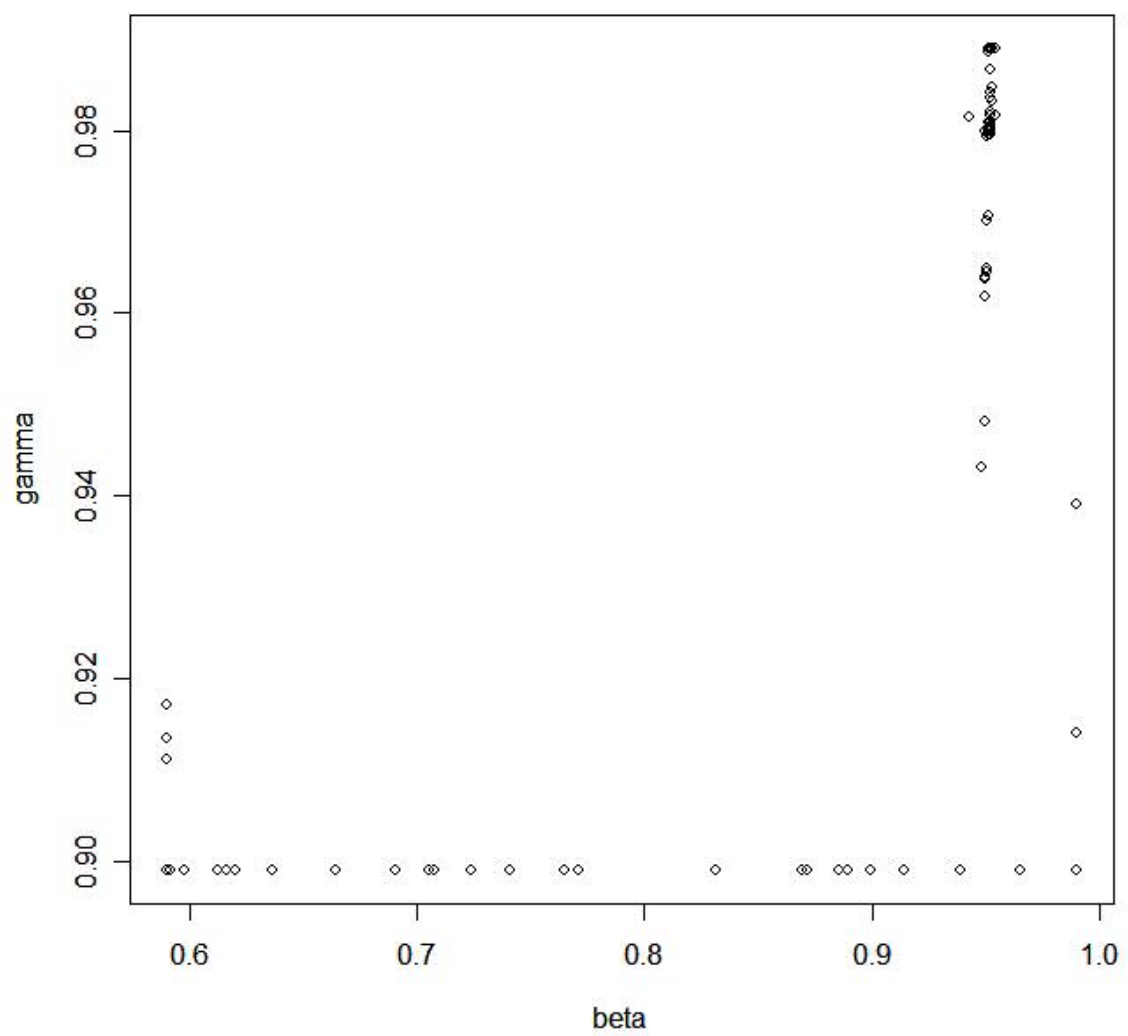46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

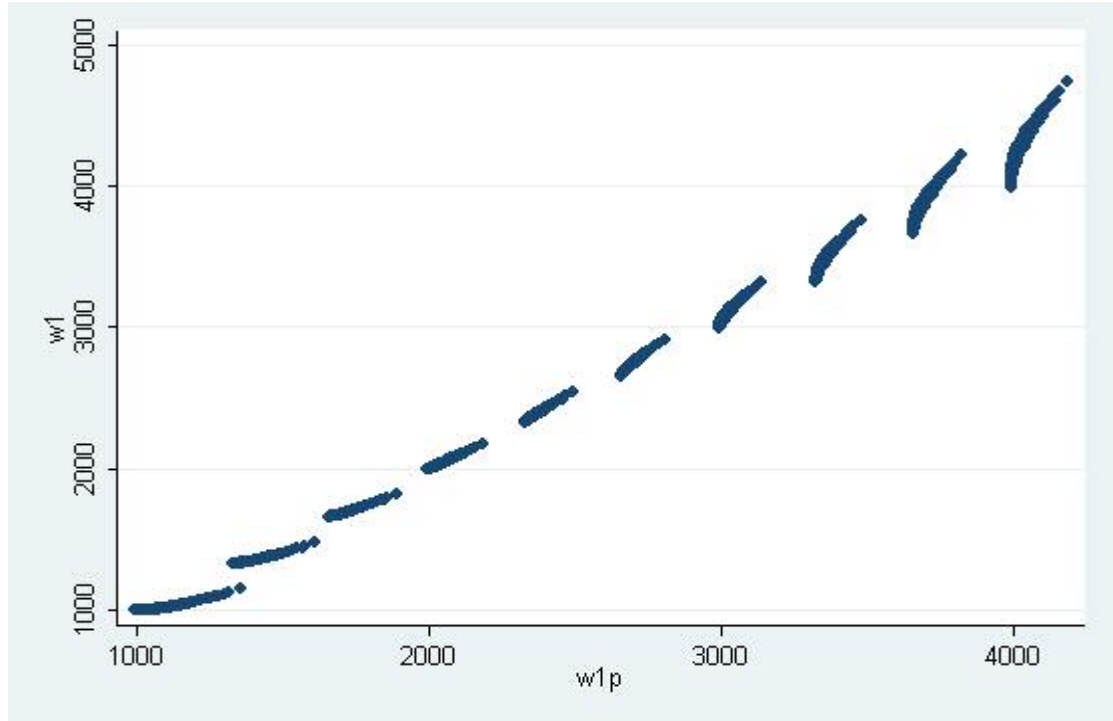
Figure 1: Sampling Distribution of Behavioral Parameters (β, γ)

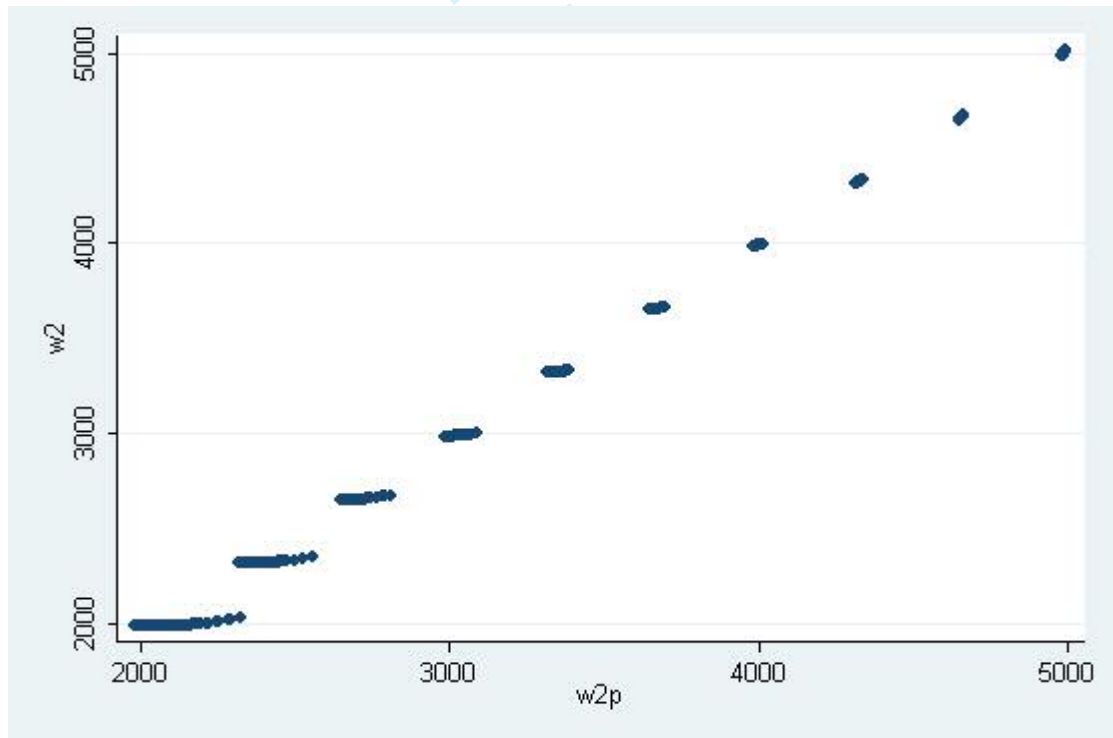Figure 2: Investment thresholds: $\bar{w}_1$ (vertically) against $\bar{w}_1^p$ (horizontally)

Figure 3: Investment thresholds: $\bar{w}_2$ (vertically) against $\bar{w}_2^p$ (horizontally)

23