



HAL
open science

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

Chris Elbers, Jan Willem Gunning, Lei Pan

► **To cite this version:**

Chris Elbers, Jan Willem Gunning, Lei Pan. Insurance and Rural Welfare: What Can Panel Data Tell Us?. Applied Economics, 2009, 41 (24), pp.3093-3101. 10.1080/00036840701367614 . hal-00582156

HAL Id: hal-00582156

<https://hal.science/hal-00582156>

Submitted on 1 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-06-0416.R1
Journal Selection:	Applied Economics
JEL Code:	C15 - Statistical Simulation Methods Monte Carlo Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, D13 - Household Production and Intrahousehold Allocation < D1 - Household Behavior and Family Economics < D - Microeconomics, C61 - Optimization Techniques Programming Models Dynamic Analysis < C6 - Mathematical Methods and Programming < C - Mathematical and Quantitative Methods, C23 - Models with Panel Data < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods
Keywords:	Insurance and Risk, Household dynamics, Structural model estimation



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

1
2
3 distribution of shocks is known. Alternatively, if locations differ in risk exposure this
4
5 can be exploited as a natural experiment to infer how behavior would change
6
7 under insurance. Clearly, the villages must then be similar in all other respects.
8
9 Matching or differencing offers some scope for relaxing this requirement but only if
10
11 all relevant variables are observed. In practice unobserved heterogeneity is likely
12
13 to be a major problem. In that case there is no alternative to estimating a structural
14
15 model, using panel data. Such a model can be used to derive how agents would
16
17 behave if insurance were available, even if no agent had been observed in that
18
19 situation. Obviously, this requires estimation of the model's structural parameters.
20
21 An important advantage of this procedure is that the researcher does not need to
22
23 know the distribution of the shocks to which the households are exposed;
24
25 estimates of the parameters characterizing that distribution will be generated as
26
27 part of the estimation procedure.
28
29

30
31
32
33
34 Lucas (2003) suggested on the basis of a back-of-envelope calculation that
35
36 insurance could not have a substantial effect on growth. However, he considered a
37
38 situation with much less risk than is common in many developing countries.
39
40 Unfortunately, for developing countries there are few empirical studies.
41
42

43
44 A notable exception is the famous paper by Rosenzweig and Wolpin (1993)
45
46 (henceforth RW). They estimated a structural model using the ICRISAT data
47
48 collected in three Indian villages. The model describes investment behavior under
49
50 risk where investment in bullocks is the key decision variable. Bullock ownership
51
52 can take three values: 0, 1 or 2. The model is estimated from data on production
53
54 and bullock ownership, but not on consumption. RW stressed that the villagers in
55
56
57
58
59
60

1
2
3 their sample participated in an informal insurance arrangement which established a
4 floor under their consumption level. They found that the introduction of actuarially
5 fair insurance (AFI) would not be welfare improving: households were already
6 sufficiently protected through informal insurance. (Strictly speaking RW could draw
7 no such conclusion. In their analysis a household that would switch from the
8 informal arrangement to AFI would still have to pay the implicit premium of the
9 informal arrangement. Clearly, this biases the comparison in favor of the *status*
10 *quo.*) Rosenzweig and Binswanger (1993) using the same data reached the
11 opposite conclusion: the poorer ICRISAT farmers would benefit greatly from the
12 introduction of insurance. Average profits in the bottom wealth quartile would
13 increase by about one third for a reduction in weather risk of one standard
14 deviation. However, their conclusion was not based on a dynamic model: a
15 household's total wealth was taken as given. Elbers, Gunning and Kinsey (2005)
16 estimated a structural model for smallholder households in Zimbabwe. They found
17 a massive effect of the introduction of AFI: on average in their sample households
18 would accumulate a capital stock (cattle) twice as large (over a 50-year period) as
19 in the absence of insurance (when consumption smoothing is the only risk coping
20 strategy available to them). As in RW this model is estimated on the basis of
21 production and capital stock data but in this study the capital stock is a continuous
22 variable.

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
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

1
2
3 remembered. However, this advantage may well come at very high econometric
4 costs, as we show in this paper. The question we investigate is whether the RW
5 research agenda is feasible in the sense that robust conclusions about the welfare
6 effects of insurance can be derived from a structural model estimated from
7 production and capital stock data if the capital stock can take only a limited number
8 of values. Our approach is to specify a simplified version of the RW model; to use
9 the model as data generating mechanism; to find out how accurately we can
10 recover the underlying behavioral parameters by estimation, given the values of all
11 other parameters and given the true model specification; and, finally, to use the
12 findings to assess the robustness of the policy conclusion on the desirability of the
13 introduction of insurance.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

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

2 Limited Dependent Variables

Consider the following deterministic intertemporal optimization problem

$$\max_{c_t, k_t} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the constraints

$$c_t + k_{t+1} = f(k_t)$$

$$c_t, k_t \geq 0$$

$$k_0 \text{ given,}$$

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 $\hat{k}_t, t=1, \dots$ 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 \bar{w}_i and corresponding asset levels a_i so that

$$w_0 = a_0 = 0$$

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

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 \bar{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 \bar{w}_1 and \bar{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 \bar{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 be estimated.

1
2
3 RW use both methods. They assume a finite decision horizon, leading to
4
5 heterogeneity in asset holding policies across decision makers of different age.
6
7 Irreversibility of the installation of a pump leads to a shift in the production function
8
9 for pump owners⁵.

10
11
12 Whether heterogeneity can indeed be used to solve the estimation problem is an
13
14 open question.⁶ RW in fact fix the value of one of the three behavioral parameters,
15
16 the discount factor. Further, as we will show for a simplified version of their model,
17
18 but with similar parameter values, heterogeneity introduced by pump ownership
19
20 does little to improve the situation since conditionally on pump-less threshold
21
22 values, behavioral parameters have almost no effect on pump-inclusive thresholds.
23
24 Also, age-based heterogeneity has very little effect on investment behavior except
25
26 for households with elderly heads; this gives such households inordinate weight in
27
28 the estimation procedure.
29
30
31
32

33
34 We conclude that estimation of behavioral parameters underlying an accumulation
35
36 process is inherently difficult if the asset involved can take (or is observed to take)
37
38 only a few discrete values. The problem can be solved by putting additional
39
40 constraints on the parameters, but any counterfactual analyses based on the
41
42 estimated parameters must then be checked for robustness against such
43
44 constraints. We also suggest that the identification problem can be solved if the
45
46 asset becomes continuous.
47
48
49
50

51
52 ⁵ We do not measure the effect of irreversibility in this paper but only keep this assumption to be in
53
54 line with RW.

55 ⁶ Rust (1994) provides detailed discussion of the identification of dynamic discrete choice models,
56
57 presenting sufficient conditions under which the model can be identified and a simple and general
58
59 estimation theory is available. One of the restrictions required in his discussion is that shocks enter
60
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:

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

subject to:

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

k_0 is given

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

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

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

1
2
3 before any support from the insurance arrangement, w is wealth at hand (available
4 for consumption and investment), k is the capital stock (constrained to take the
5 values 0, 1 or 2) with price p_k , P is a dummy variable indicating whether the
6 household owns a pump (with price p_p) and ε is an income shock (a draw from a
7 distribution known by the agent). Income shocks are independent over time. The
8 function $\alpha(k)$ takes the values $\alpha_0, \alpha_1, \alpha_2$ for $k = 0, 1, 2$ respectively; this picks up
9 the effect of bullock ownership on income. The productivity of pump is captured by
10 α_p . The age of the household head (which determines the time remaining until
11 time T) is denoted by τ . In this formulation investment in pumps (which is
12 irreversible) is exogenous (contrary to the RW formulation). Agents without a pump
13 do not expect to acquire one. Expected utility is maximized with respect to the
14 policy function φ_τ .

15
16 Note from (1) and (3) that the insurance arrangement will pay out only to
17 households without cattle: a household with $c < c_{\min}$ will have to sell its cattle before
18 it is entitled to consumption support since c is defined as the difference between
19 wealth at hand (w) which includes the cattle asset and the investment.

20
21 The optimal policy function will depend on τ and P . Given these two parameters
22 the function will involve two threshold values for wealth at hand, \bar{w}_1, \bar{w}_2 such that in

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 arrangement. It would be more natural to specify different parameters for these roles so that the
58 welfare effect of a change in the protection offered by the insurance scheme can be analyzed.
59
60

each period the household chooses $k = 1$ for $\bar{w}_1 \leq w < \bar{w}_2$, $k = 2$ for $w \geq \bar{w}_2$ and $k = 0$ otherwise⁸.

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 \bar{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 \bar{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 $\bar{w}_1 = 2461$ and $\bar{w}_2 = 3453$ 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 $\bar{w}_1^P = 2461$ and $\bar{w}_2^P = 3453$ 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 eight¹⁰ 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 \tilde{w}_t . The dataset now consists of a vector $\{k_1, \tilde{w}_1, \dots, k_s, \tilde{w}_s \mid k_0, P_t\}$ 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

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

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 distribution¹² $N((\varepsilon + \eta)\sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_\eta^2), \sigma_\varepsilon^2 \sigma_\eta^2 / \sqrt{\sigma_\varepsilon^2 + \sigma_\eta^2})$ with cumulative distribution function Γ . For instance, if the household has no pump

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

Note that the likelihood depends only indirectly, through the threshold values \bar{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 interpolation¹³ 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

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

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

1
2
3 the estimates of β, γ and c_{\min} should be close to their true values if they could be
4
5
6 estimated from the model the choice of different bounds do not make difference to
7
8
9 the estimation of the parameters.

10 11 **3.3 Results**

12
13
14
15
16
17 With the above choice of model and parameters it turns out that the threshold
18
19 values $\bar{w}_{1,2}$ are virtually insensitive to parameters β and γ . Hence it is impossible
20
21
22 to estimate these parameters with any accuracy. Given the typical size of
23
24 household shocks and only 30 households, the influence of β and γ on the
25
26 likelihood function is almost absent. We find that γ tends to settle on one of the
27
28 bounds and β tends to stay very close to its initial value. Figure 2 is a typical
29
30 scatter plot of the joint (β, γ) sampling distribution, based on 100 simulations.
31
32

33
34
35 On the other hand, c_{\min} can be estimated fairly accurately. Figure 3 displays the
36
37 sampling distribution of the estimator for c_{\min} . The mean and standard deviation
38
39 determined from 100 simulations are 1444.7 and 80.3.
40
41

42
43 Recall that the heterogeneity in pump ownership raises the number of thresholds
44
45 above the number of parameters so that the problem discussed in section 2 of
46
47 having to recover three coefficients from two threshold values does not arise.
48
49 However, it turns out that the extra pair of thresholds do not convey sufficient
50
51 additional information: as may be seen from Figures 4 and 5 the thresholds for
52
53 pump-owning households are almost insensitive to the behavioral parameters
54
55 given the corresponding thresholds for the other households. Hence the second
56
57
58
59
60

1
2
3 pair of thresholds convey little extra information on the parameters. In fact, we find
4
5 that the thresholds are almost exclusively determined by the value of c_{\min} , so that
6
7 the combined figures trace essentially a one-dimensional sub-set in the space of
8
9 threshold values. This is why only a single parameter can be estimated with
10
11 reasonable precision.
12
13

14
15 What are the implications for the robustness of the conclusion that actuarially fair
16
17 insurance would not raise welfare? We investigate this by calculating ΔW , the net
18
19 increase in welfare (i.e. the expected value of discounted utility) as a result of
20
21 replacing the informal arrangement supporting c_{\min} with AFI.¹⁴ We do this for each
22
23 of six household types, defined by whether they have (initially) a pump and
24
25 whether they start with 0, 1 or 2 bullocks. Table 1 shows ΔW for each of these
26
27 household types and for various values of β in the (0.59, 0.99) range, and putting
28
29 all other parameters to their true value.¹⁵ The values shown in the Table are
30
31 calculated under the assumption that no household will acquire a pump: the first
32
33 three household types remain pumpless throughout.
34
35
36
37
38

39
40 The Table shows that the first two types of households (those who initially have no
41
42 pump and at most one bullock) would not be willing to give up the informal
43
44 insurance arrangement (which guarantees a consumption level c_{\min}) in exchange
45
46

47
48 ¹⁴ Recall that this measure is biased since (as RW recognise) the cost of the informal arrangement
49
50 would be reflected in a premium which would not have to be paid under AFI. Since RW do not know
51
52 this cost (which is reflected in $\alpha(k)$) they ignore it in the comparison: hence under AFI the
53
54 household continues to pay the same premium. This is a major issue: given the choice between AFI
55
56 (i.e. constant consumption) and an arrangement where the household would receive positive
57
58 shocks but (as a result of the c_{\min} floor) no negative shocks it might well prefer the latter if the two
59
60 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.

1
2
3 for AFI. The reason is simple: under AFI they would no longer enjoy the positive
4 income shocks which they experience under the (asymmetric) informal insurance
5 arrangement. For richer households this advantage of the c_{\min} arrangement is
6 offset by a larger difference between c_{\min} and mean consumption: the informal
7 arrangement gives them little downward protection. As a result, households with
8 pumps or with two bullocks would switch to AFI. Clearly, the aggregate effect
9 depends on the distribution of the population over household types and on the
10 discount factor. In the Table we show a particular distribution. Under this
11 distribution the change in aggregate welfare depends on the value of the discount
12 factor. For low values of β (i.e. a high discount rate) the net effect is negative, for
13 values of 0.79 or higher it is positive: AFI would be accepted.

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30 This example underestimates the case for AFI, for two reasons. First, as noted
31 above, under AFI the implicit premium of the informal insurance would no longer
32 have to be paid but this is not taken into account. Secondly, we have treated pump
33 investment as exogenous. In the RW world pump investment is endogenous. A
34 pumpless household might receive (at some future date) a positive shock large
35 enough to enable it to buy a pump. At that stage insurance would become
36 attractive. Therefore, if households were not forced to adopt insurance now or stay
37 with the informal arrangement forever, but were instead offered the *option* of
38 switching to insurance welfare then gains might be positive even for the first two
39 types of households.

40
41
42
43
44
45
46
47
48
49
50
51
52
53
54 Since β cannot be estimated with precision, Table 1 implies that any policy
55 conclusion on the welfare effect of introducing formal insurance will not be robust.
56
57
58
59
60

1
2
3 If β is estimated (in a relatively small sample) it may easily settle on one of the two
4
5 boundary values (with opposite policy conclusions). Conversely, if β is fixed (as in
6
7
8
9 RW, who set $\beta = 0.95$) the policy conclusion is thereby fixed as well.

10 11 12 13 14 15 16 17 **4 Conclusion**

18
19
20
21
22 There is a renewed interest in insurance mechanisms to assist rural households in
23
24 risk coping. It is rarely possible to evaluate such interventions through (quasi)
25
26 experimental evaluation methods. In principle estimating a structural model of
27
28 household behavior under risk (using panel data) is a viable alternative.
29
30 Rosenzweig and Wolpin (1993) applied this approach to the ICRISAT data and
31
32 found that insurance would not be welfare improving. In this paper we have
33
34 investigated the robustness of this policy conclusion. We have considered a class
35
36 of models of household behavior under risk where assets can take only a small
37
38 number of values. This severely restricts the scope for estimating structural
39
40 coefficients. In the RW case heterogeneity could solve this problem but we have
41
42 shown (for a simplified version of their model) that in small samples two of the
43
44 coefficients cannot be estimated with reasonable precision in spite of this
45
46 heterogeneity. Since the policy conclusion on the desirability of introducing formal
47
48 insurance is sensitive to the value of these coefficients, the conclusion is probably
49
50 not robust. This does not mean that we cannot use panel data to assess the scope
51
52 for insurance. Rather, it implies that if asset data indeed take only a small number
53
54
55
56
57
58
59
60

1
2
3 of values (relative to the number of parameters to be estimated) then estimation
4
5 requires heterogeneity which (unlike the heterogeneity allowed for by RW) leads to
6
7 *independent* variation in threshold values. The procedure we have described can
8
9 easily establish whether this condition is satisfied. Another way to solve the
10
11 problem is by treating asset variable(s) continuous, e.g., by using livestock (an
12
13 aggregate of cattle, goats, sheep etc.) rather than bullocks as the asset variable.
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

For Peer Review

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.

1
2
3
4
5
6 Rosenzweig, Mark R. and Kenneth Wolpin (1993) Credit Market Constraints,
7
8 Consumption Smoothing, and the Accumulation of Durable Production Assets in
9
10 Low-Income Countries: Investment in Bullocks in India, *Journal of Political*
11
12 *Economy*, vol.101, pp. 223-244.
13
14

15
16
17 Rust (1994) Structural estimation of Markov decision processes, in Handbook of
18
19 Econometrics, vol. 4, ed. by R. Engle and D. McFadden, Amsterdam: North
20
21 Holland.
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

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

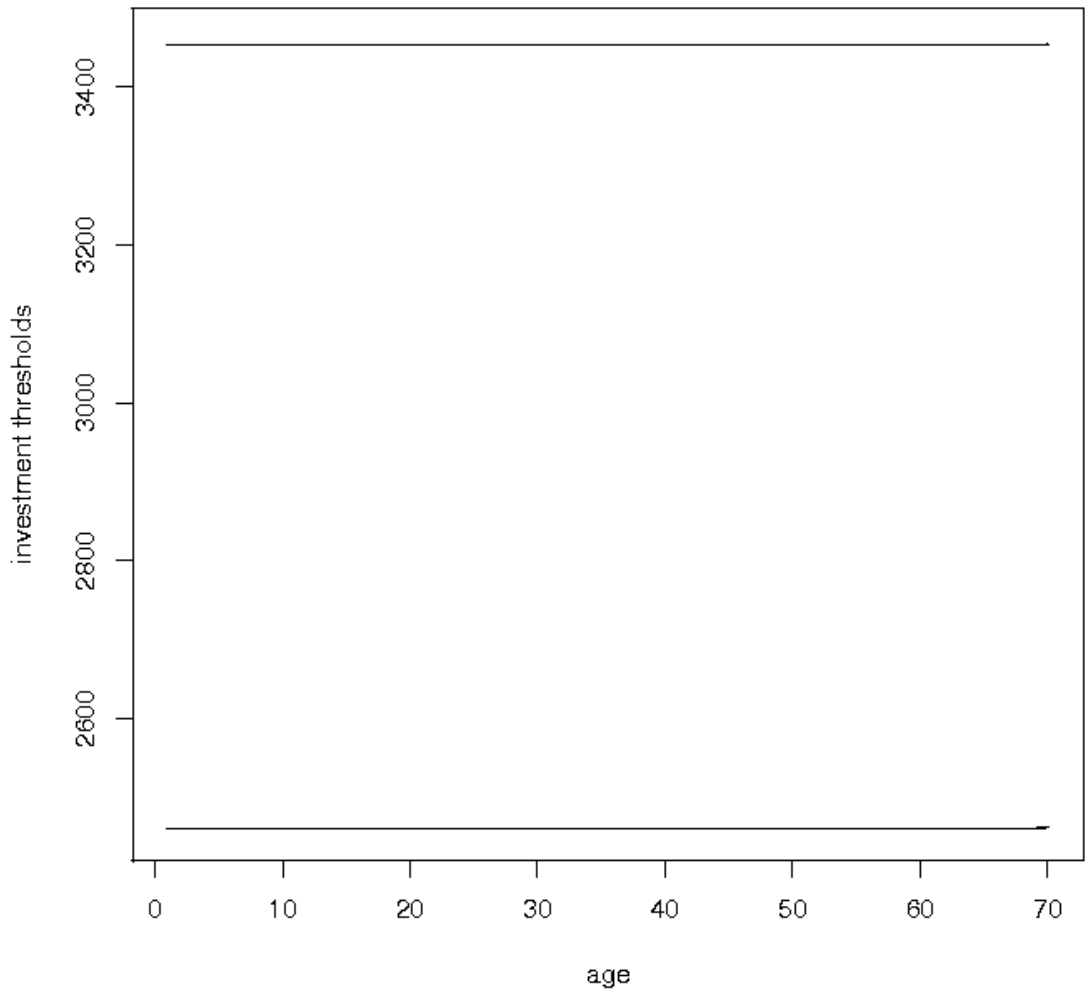


Figure 1: Investment thresholds \bar{w}_1 (below) and \bar{w}_2 (above) for households without a pump.

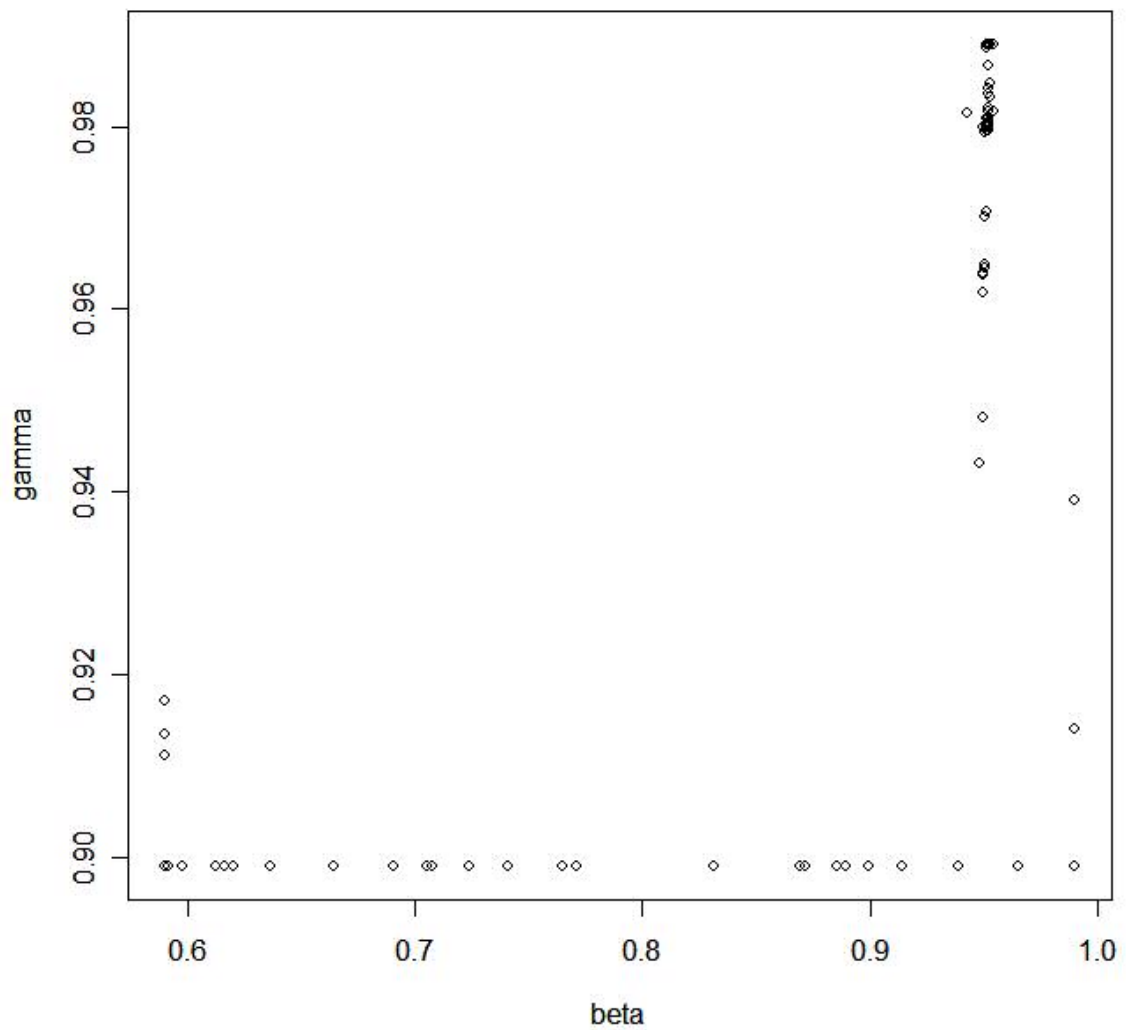


Figure 2: Sampling Distribution of Behavioral Parameters (β , γ)

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

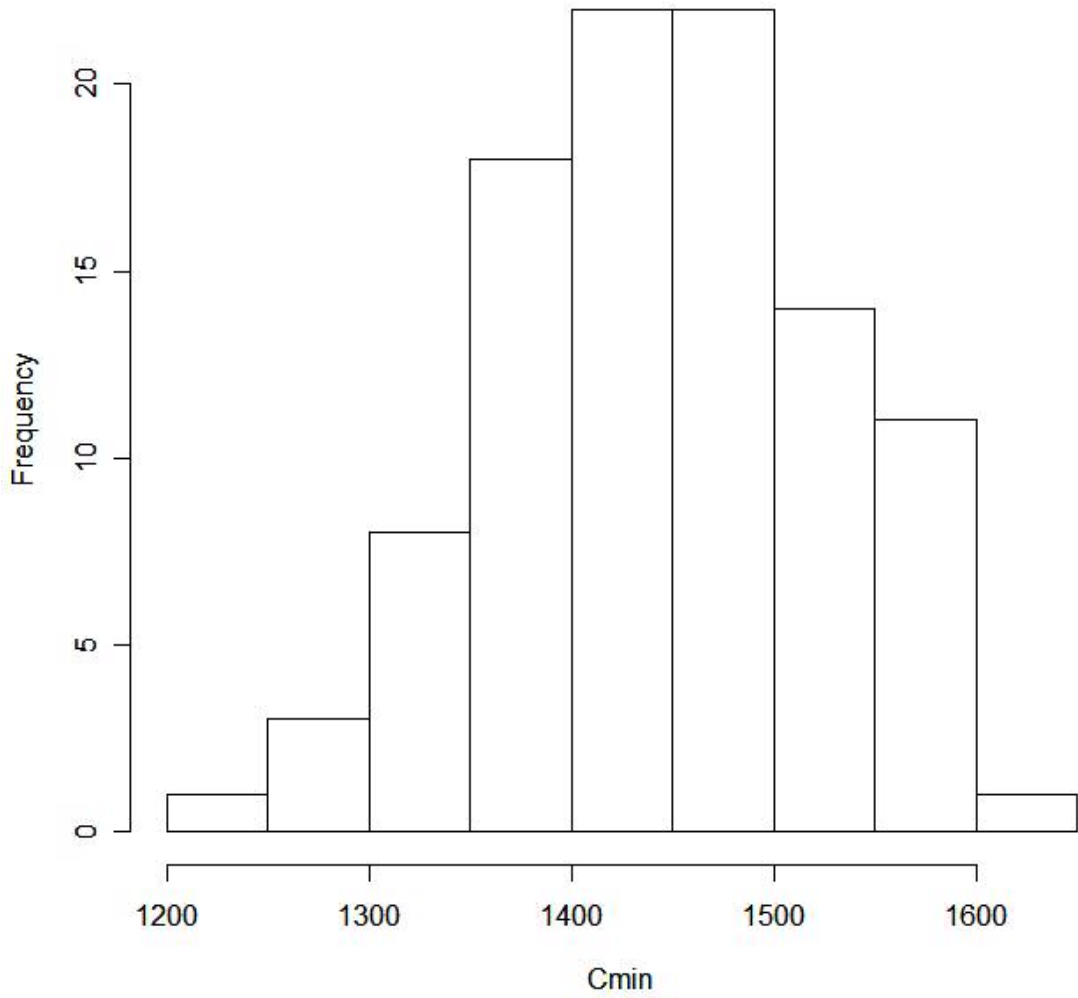


Figure 3: Sampling Distribution of the c_{\min} Estimator

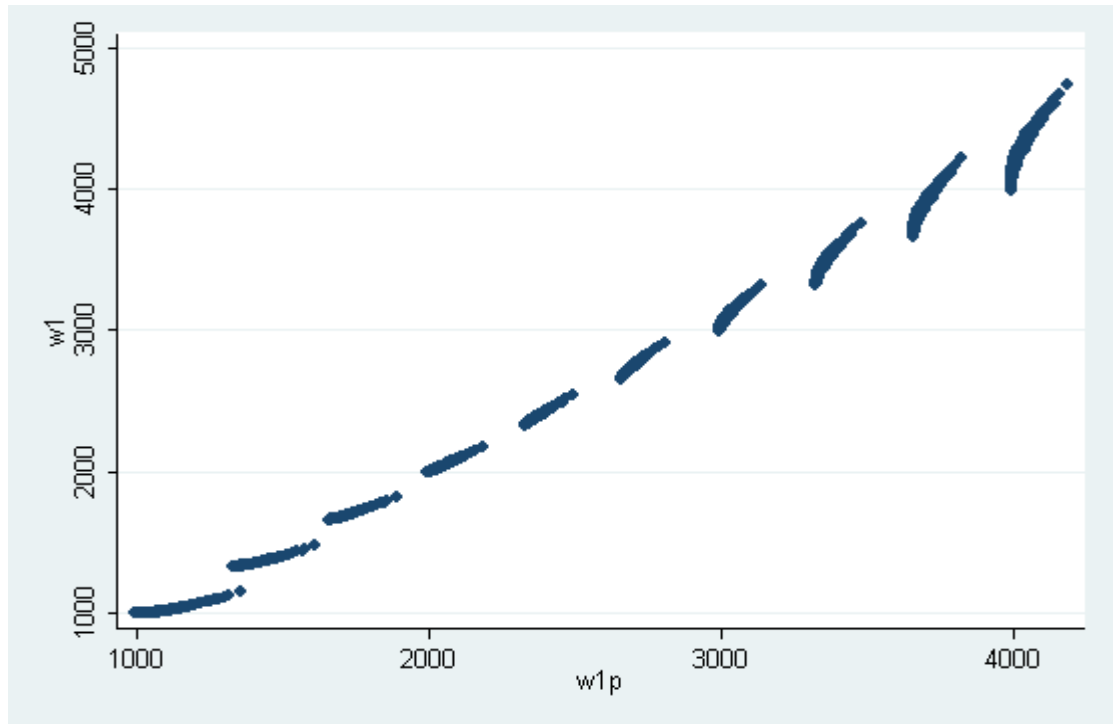


Figure 4: Investment thresholds: \bar{w}_1 (vertically) against \bar{w}_1^p (horizontally)

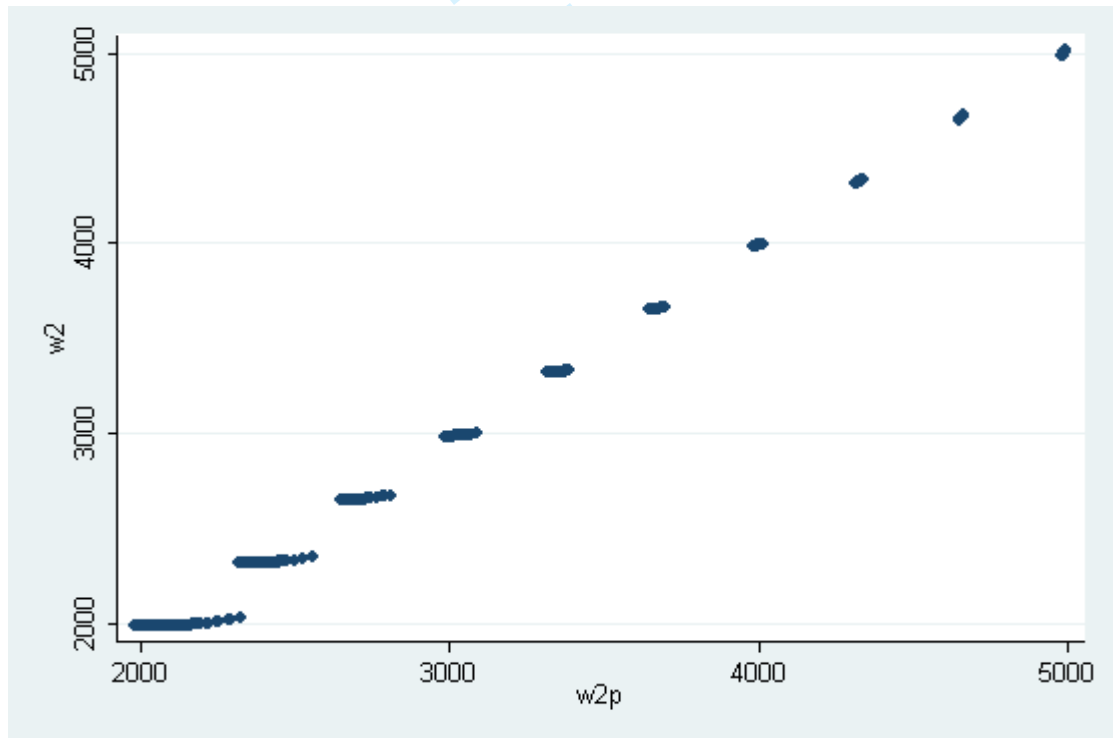


Figure 5: Investment thresholds: \bar{w}_2 (vertically) against \bar{w}_2^p (horizontally)

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
2
3 risk.¹ Here the implicit premium is the loss of production as a result of using assets
4
5 for this purpose.
6

7
8 Ideally, the scope for insurance can be evaluated by comparing villages with and
9
10 without insurance where treatment villages have been selected randomly and the
11
12 distribution of shocks is known. Alternatively, if locations differ in risk exposure this
13
14 can be exploited as a natural experiment to infer how behavior would change
15
16 under insurance. Clearly, the villages must then be similar in all other respects.
17
18 Matching or differencing offers some scope for relaxing this requirement but only if
19
20 all relevant variables are observed. In practice unobserved heterogeneity is likely
21
22 to be a major problem. In that case there is no alternative to estimating a structural
23
24 model, using panel data. Such a model can be used to derive how agents would
25
26 behave if insurance were available, even if no agent had been observed in that
27
28 situation. Obviously, this requires estimation of the model's structural parameters.
29
30 An important advantage of this procedure is that the researcher does not need to
31
32 know the distribution of the shocks to which the households are exposed;
33
34 estimates of the parameters characterizing that distribution will be generated as
35
36 part of the estimation procedure.
37
38
39
40
41
42

43 Lucas (2003) suggested on the basis of a back-of-envelope calculation that
44
45 insurance could not have a substantial effect on growth. However, he considered a
46
47 situation with much less risk than is common in many developing countries.
48
49 Unfortunately, for developing countries there are few empirical studies.
50
51

52 A notable exception is the famous paper by Rosenzweig and Wolpin (1993)
53
54 (henceforth RW). They estimated a structural model using the ICRISAT data
55
56

57
58 ¹ See e.g. Tanner (1997) for US evidence and Elbers *et al.* (forthcoming) for Zimbabwe evidence.
59
60

1
2
3 collected in three Indian villages. The model describes investment behavior under
4 risk where investment in bullocks is the key decision variable. Bullock ownership
5 can take three values: 0, 1 or 2. The model is estimated from data on production
6 and bullock ownership, but not on consumption. RW stressed that the villagers in
7 their sample participated in an informal insurance arrangement which established a
8 floor under their consumption level. They found that the introduction of actuarially
9 fair insurance (AFI) would not be welfare improving: households were already
10 sufficiently protected through informal insurance. Rosenzweig and Binswanger
11 (1993) using the same data reached the opposite conclusion: the poorer ICRISAT
12 farmers would benefit greatly from the introduction of insurance. Average profits in
13 the bottom wealth quartile would increase by about one third for a reduction in
14 weather risk of one standard deviation. However, their conclusion was not based
15 on a dynamic model: a household's total wealth was taken as given. Elbers,
16 Gunning and Kinsey (forthcoming) estimate a structural model for smallholder
17 households in Zimbabwe. They found a massive effect of the introduction of AFI:
18 on average in their sample households would accumulate a capital stock (cattle)
19 twice as large (over a 50-year period) as in the absence of insurance (when
20 consumption smoothing is the only risk coping strategy available to them). As in
21 RW this model is estimated on the basis of production and capital stock data but in
22 this study the capital stock is a continuous variable.

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
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

1
2
3 quite accurately. An implication is that observing the number of cattle is likely to be
4
5 less subject to measurement error than, say, consumption. However, this
6
7 advantage may well come at very high econometric costs, as we show in this
8
9 paper. The question we investigate is whether the RW research agenda is feasible
10
11 in the sense that robust conclusions about the welfare effects of insurance can be
12
13 derived from a structural model estimated from production and capital stock data if
14
15 the capital stock can take only a limited number of values. Our approach is to
16
17 specify a simplified version of the RW model; to use the model as data generating
18
19 mechanism; to find out how accurately we can recover the underlying behavioral
20
21 parameters by estimation, given the values of all other parameters and given the
22
23 true model specification; and, finally, to use the findings to assess the robustness
24
25 of the policy conclusion on the desirability of the introduction of insurance.
26
27
28
29
30

31
32 The structure of the paper is as follows. In the next section we show under what
33
34 conditions the use of limited dependent variables in this class of models may lead
35
36 to large standard errors. In section 3 we specify the model and use simulation
37
38 experiments to derive the distribution of the structural coefficients. We find, as
39
40 expected, very large errors. It turns out that in this class of models the RW
41
42 conclusion as to the welfare effects of insurance is not robust. Section 4 concludes.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

2 Limited Dependent Variables

Consider the following deterministic intertemporal optimization problem

$$\max_{c_t, k_t} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the constraints

$$c_t + k_{t+1} = f(k_t)$$

$$c_t, k_t \geq 0$$

$$k_0 \text{ given,}$$

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 $\hat{k}_t, t=1, \dots$ 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).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 \bar{w}_i and corresponding asset levels a_i so that

$$w_0 = a_0 = 0$$

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

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 \bar{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 \bar{w}_1 and \bar{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 \bar{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 be estimated.

1
2
3 RW use both methods. They assume a finite decision horizon, leading to
4
5 heterogeneity in asset holding policies across decision makers of different age.
6
7 Irreversibility of the installation of a pump leads to a shift in the production function
8
9 for pump owners.
10

11
12 Whether heterogeneity can indeed be used to solve the estimation problem is an
13
14 open question.⁶ RW in fact fix the value of one of the three behavioral parameters,
15
16 the discount factor. Further, as we will show for a simplified version of their model,
17
18 but with similar parameter values, heterogeneity introduced by pump ownership
19
20 does little to improve the situation since conditionally on pump-less threshold
21
22 values, behavioral parameters have almost no effect on pump-inclusive thresholds.
23
24 Also, age-based heterogeneity has very little effect on investment behavior except
25
26 for households with elderly heads; this gives such households inordinate weight in
27
28 the estimation procedure.
29
30
31
32

33
34 We conclude that estimation of behavioral parameters underlying an accumulation
35
36 process is inherently difficult if the asset involved can take only a few discrete
37
38 values. The problem can be solved by putting additional constraints on the
39
40 parameters, but any counterfactual analyses based on the estimated parameters
41
42 must then be checked for robustness against such constraints. We also suggest
43
44 that the identification problem can be solved if the asset becomes continuous.
45
46
47
48
49
50
51
52

53
54 ⁶ Rust (1994) provides detailed discussion of the identification of dynamic discrete choice models,
55
56 presenting sufficient conditions under which the model can be identified and a simple and general
57
58 estimation theory is available. One of the restrictions required in his discussion is that shocks enter
59
60 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:

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

subject to:

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

k_0 is given

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

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

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.

1
2
3 before any support from the insurance arrangement, w is wealth at hand (available
4 for consumption and investment), k is the capital stock (constrained to take the
5 values 0, 1 or 2) with price p_k , P is a dummy variable indicating whether the
6 household owns a pump (with price p_p) and ε is an income shock (a draw from a
7 distribution known by the agent). Income shocks are independent over time. The
8 function $\alpha(k)$ takes the values $\alpha_0, \alpha_1, \alpha_2$ for $k = 0, 1, 2$ respectively; this picks up
9 the effect of bullock ownership on income. The productivity of pumps is captured
10 by α_p . The age of the household head (which determines the time remaining until
11 time T) is denoted by τ . In this formulation investment in pumps (which is
12 irreversible) is exogenous (contrary to the RW formulation). Agents without a pump
13 do not expect to acquire one. Expected utility is maximized with respect to the
14 policy function φ_τ .

15
16 Note from (1) and (3) that the insurance arrangement will pay out only to
17 households without cattle: a household with $c < c_{\min}$ will have to sell its cattle before
18 it is entitled to consumption support since c is defined as the difference between
19 wealth at hand (w) which includes the cattle asset and the investment.

20
21 The optimal policy function will depend on τ and P . Given these two parameters
22 the function will involve two threshold values for wealth at hand, \bar{w}_1, \bar{w}_2 such that in
23 each period the household chooses $k = 1$ for $\bar{w}_1 \leq w < \bar{w}_2$, $k = 2$ for $w \geq \bar{w}_2$ and $k = 0$
24 otherwise⁹.

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

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 \bar{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 \bar{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 $\bar{w}_1 = 2461$ and $\bar{w}_2 = 3453$ for households without a pump and again $\bar{w}_1^P = 2461$ and $\bar{w}_2^P = 3453$ 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 eight¹¹ 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 $\{k_1, \tilde{w}_1, \dots, k_s, \tilde{w}_s \mid k_0, P_t\}$ 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 distribution¹³ $N((\varepsilon + \eta)\sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_\eta^2), \sigma_\varepsilon^2 \sigma_\eta^2 / \sqrt{\sigma_\varepsilon^2 + \sigma_\eta^2})$ with cumulative distribution function Γ . For instance, if the household has no pump

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

Note that the likelihood depends only indirectly, through the threshold values \bar{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 interpolation¹⁴ 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).

1
2
3 the estimates of β, γ and c_{\min} should be close to their true values if they could be
4
5
6 estimated from the model the choice of different bounds do not make difference to
7
8
9 the estimation of the parameters.

10 11 **3.3 Results**

12
13
14
15
16
17 With the above choice of model and parameters it turns out that the threshold
18
19 values $\bar{w}_{1,2}$ are virtually insensitive to parameters β and γ . Hence it is impossible
20
21
22 to estimate these parameters with any accuracy. We find that γ tends to settle on
23
24
25 one of the bounds and β tends to stay very close to its initial value. Figure 1 is a
26
27
28 typical scatter plot of the joint (β, γ) sampling distribution, based on 100
29
30
31 simulations.

32
33 On the other hand, c_{\min} can be estimated fairly accurately. The mean and standard
34
35
36 deviation determined from 100 simulations are 1444.7 and 80.3.

37
38 Recall that the heterogeneity in pump ownership raises the number of thresholds
39
40
41 above the number of parameters so that the problem discussed in section 2 of
42
43
44 having to recover three coefficients from two threshold values does not arise.
45
46
47 However, it turns out that the extra pair of thresholds do not convey sufficient
48
49
50 additional information: as may be seen from Figures 2 and 3 the thresholds for
51
52
53 pump-owning households are almost insensitive to the behavioral parameters
54
55
56 given the corresponding thresholds for the other households. Hence the second
57
58
59 pair of thresholds convey little extra information on the parameters. In fact, we find
60
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.

1
2
3 arrangement. For richer households this advantage of the c_{\min} arrangement is
4
5
6 offset by a larger difference between c_{\min} and mean consumption: the informal
7
8 arrangement gives them little downward protection. As a result, households with
9
10 pumps or with two bullocks would switch to AFI. Clearly, the aggregate effect
11
12 depends on the distribution of the population over household types and on the
13
14 discount factor. In the Table we show a particular distribution. Under this
15
16 distribution the change in aggregate welfare depends on the value of the discount
17
18 factor. For low values of β (i.e. a high discount rate) the net effect is negative, for
19
20 values of 0.79 or higher it is positive: AFI would be accepted.
21
22
23

24
25 This example underestimates the case for AFI, for two reasons. First, as noted
26
27 above, under AFI the implicit premium of the informal insurance would no longer
28
29 have to be paid but this is not taken into account. Secondly, we have treated pump
30
31 investment as exogenous. In the RW world pump investment is endogenous. A
32
33 pumpless household might receive (at some future date) a positive shock large
34
35 enough to enable it to buy a pump. At that stage insurance would become
36
37 attractive. Therefore, if households were not forced to adopt insurance now or stay
38
39 with the informal arrangement forever, but were instead offered the *option* of
40
41 switching to insurance then welfare gains might be positive even for the first two
42
43 types of households.
44
45
46
47

48
49 Since β cannot be estimated with precision, Table 1 implies that any policy
50
51 conclusion on the welfare effect of introducing formal insurance will not be robust.
52
53 If β is estimated (in a relatively small sample) it may easily settle on one of the two
54
55
56
57
58
59
60

1
2
3 boundary values (with opposite policy conclusions). Conversely, if β is fixed (as in
4
5
6 RW, who set $\beta = 0.95$) the policy conclusion is thereby fixed as well.
7
8
9

10 11 12 13 14 **4 Conclusion** 15 16 17

18
19 There is a renewed interest in insurance mechanisms to assist rural households in
20 risk coping. It is rarely possible to evaluate such interventions through (quasi)
21 experimental evaluation methods. In principle estimating a structural model of
22 household behavior under risk (using panel data) is a viable alternative.
23
24 Rosenzweig and Wolpin (1993) applied this approach to the ICRISAT data and
25 found that insurance would not be welfare improving. In this paper we have
26 investigated the robustness of this policy conclusion. We have considered a class
27 of models of household behavior under risk where assets can take only a small
28 number of values. This severely restricts the scope for estimating structural
29 coefficients. In the RW case heterogeneity could solve this problem but we have
30 shown (for a simplified version of their model) that in small samples two of the
31 coefficients cannot be estimated with reasonable precision in spite of this
32 heterogeneity. Since the policy conclusion on the desirability of introducing formal
33 insurance is sensitive to the value of these coefficients, the conclusion is probably
34 not robust. This does not mean that we cannot use panel data to assess the scope
35 for insurance. Rather, it implies that if asset data indeed take only a small number
36 of values (relative to the number of parameters to be estimated) then estimation
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 requires heterogeneity which (unlike the heterogeneity allowed for by RW) leads to
4
5 *independent* variation in threshold values. The procedure we have described can
6
7 easily establish whether this condition is satisfied. An alternative is to treat assets
8
9 as continuous variables, *e.g* by using livestock (an aggregate of cattle, goats,
10
11 sheep *etc.*) rather than bullocks as the capital stock.
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

For Peer Review

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.

1
2
3 Rosenzweig, Mark R. and Hans Binswanger (1993) Wealth, Weather Risk and the
4 Composition and Profitability of Agricultural Investments, *Economic Journal*, vol.
5
6 103, pp. 56-78.
7
8
9

10
11
12 Rosenzweig, Mark R. and Kenneth Wolpin (1993) Credit Market Constraints,
13 Consumption Smoothing, and the Accumulation of Durable Production Assets in
14
15 Low-Income Countries: Investment in Bullocks in India, *Journal of Political*
16
17 *Economy*, vol.101, pp. 223-244.
18
19
20
21

22
23
24 Rust, John (1994) Structural estimation of Markov decision processes, in
25
26 *Handbook of Econometrics*, vol. 4, ed. by R. Engle and D. McFadden, Amsterdam:
27
28 North Holland.
29
30
31

32
33
34 Tanner, Evan (1997) Shifts in US Savings: Long-run Asset Accumulation versus
35
36 Consumption Smoothing, *Applied Economics*, vol. 29, pp. 989-999.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Endowment							
initial k	0	1	2	0	1	2	
pump	0	0	0	1	1	1	
discount factor							weighted 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

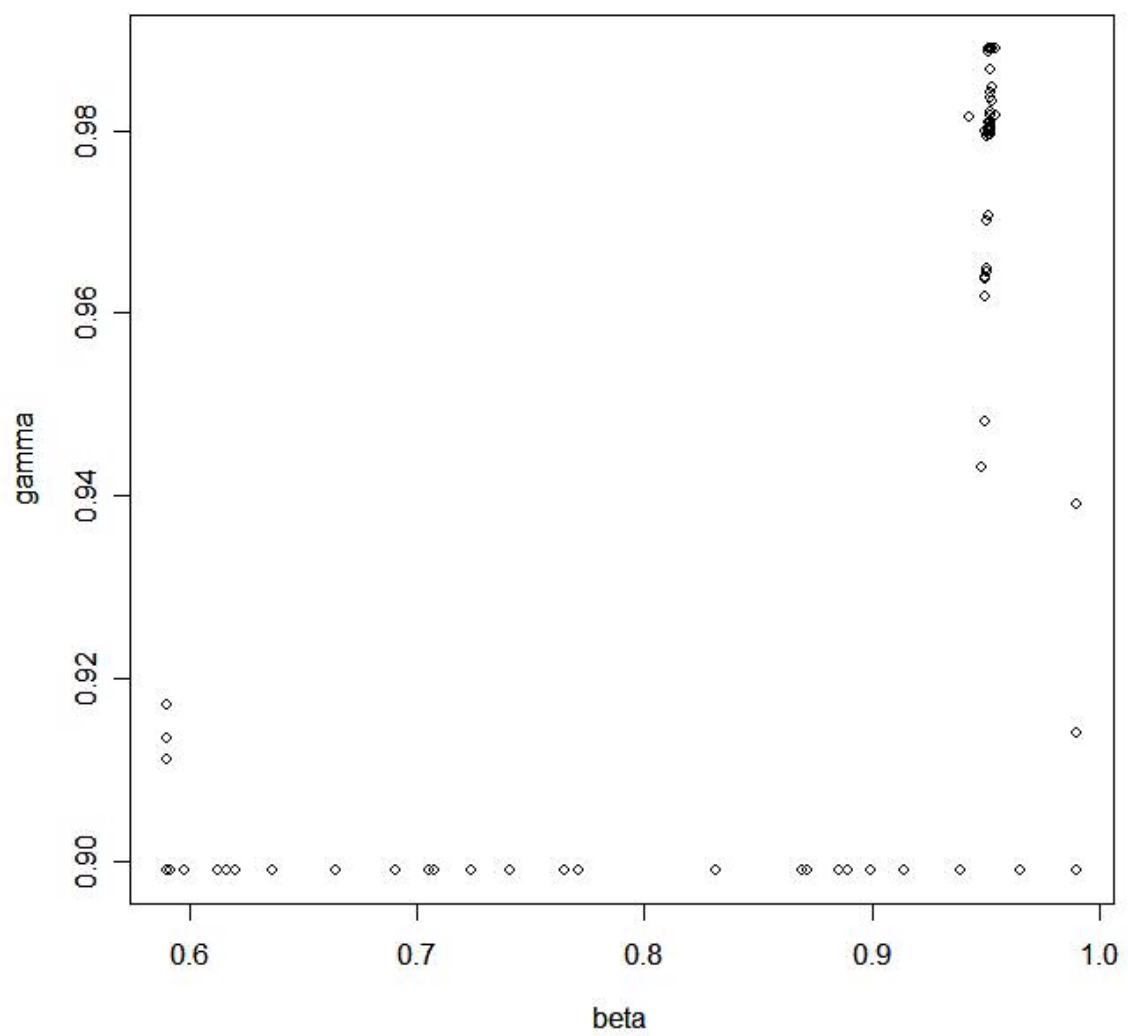


Figure 1: Sampling Distribution of Behavioral Parameters (β, γ)

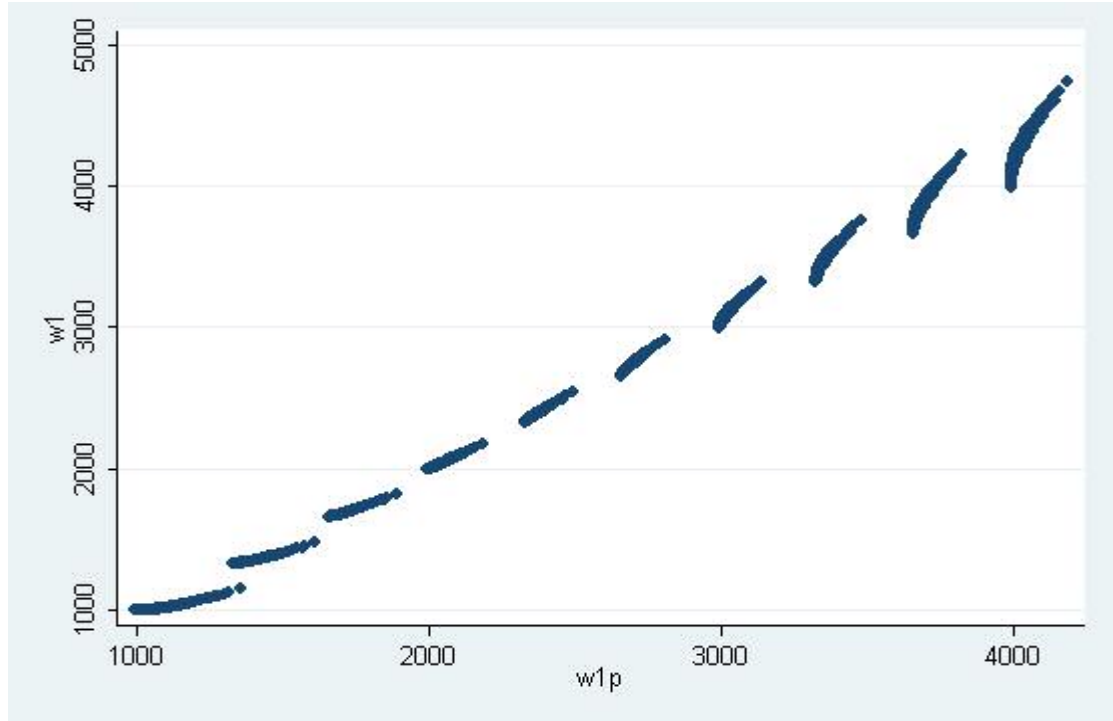


Figure 2: Investment thresholds: \bar{w}_1 (vertically) against \bar{w}_1^p (horizontally)

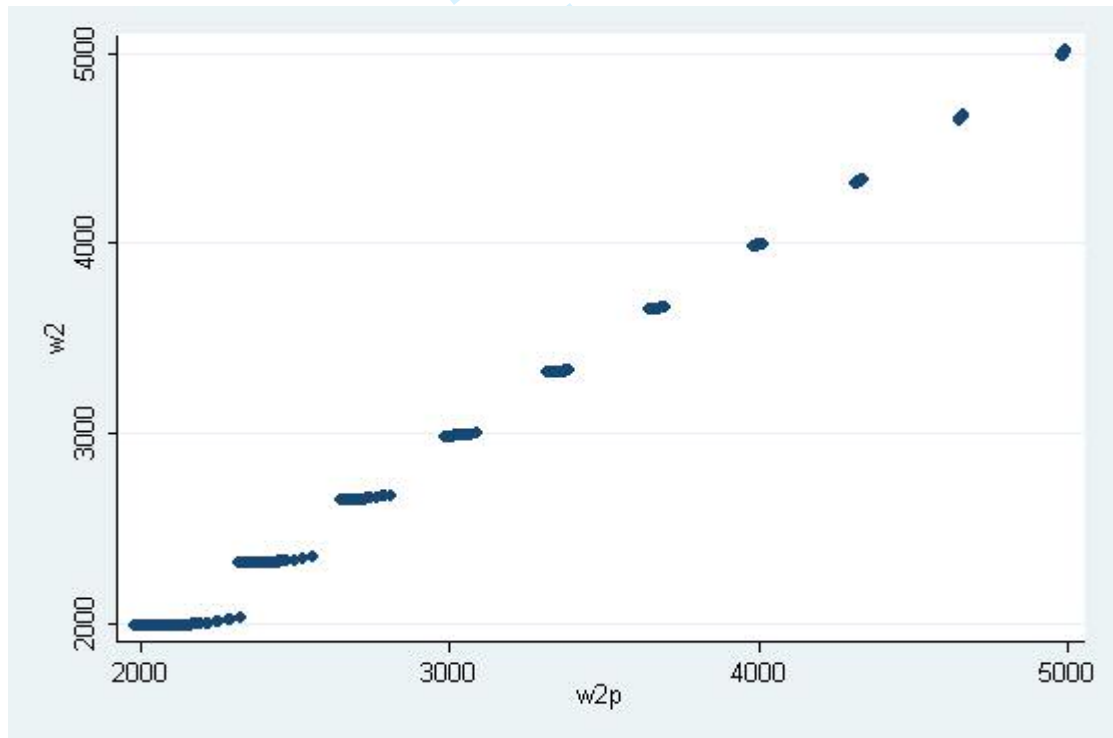


Figure 3: Investment thresholds: \bar{w}_2 (vertically) against \bar{w}_2^p (horizontally)