"Poor stays poor"

Household asset poverty traps in rural semi-arid India

Felix Naschold
Cornell University
Department of Applied Economics and Management
313a Warren Hall
Ithaca, NY 14853
fn23@cornell.edu

30 April 2009

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009
AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009

JEL classifications: I32,C14,O12
Keywords: Household Welfare Dynamics, Semiparametric Estimation, Penalized Splines, India, Panel Data, Asset Poverty
Word Count: 7998 words

I would like to thank Chris Barrett, David Ruppert, Ravi Kanbur and Per Pinstrup-Andersen for comments and guidance. An earlier version also benefited from feedback by conference participants at WIDER, the NEUDC, the AAEA and Cornell University. I would also like to thank the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) for making available the data, and the United States Agency for International Development (USAID) for financial support through grant LAG-A-00-96-90016-00 to the BASIS CRSP. Any views expressed and any remaining errors are mine alone.

Abstract

Although identifying the existence and the nature of household-level poverty traps would have important implications for the design of poverty reduction policies empirical evidence is still scant. A small, but growing empirical literature has begun testing for poverty traps as thresholds in non-linear welfare dynamics. Employing a variety of quantitative methods it has produced a variety of conclusions. This paper uses a uniquely long household panel from three villages in rural India to examine whether the detection of poverty traps may be contingent on the quantitative method used to model household welfare dynamics. It then employs a novel semiparametric panel data estimator that combines the advantages of the existing methods. Since in the context of dynamic poverty traps we are primarily concerned with expected, structural well-being it measures household welfare in assets. Structural immobility in these Indian villages is pervasive. Household asset holdings are stagnant over time. Absent any structural changes, the currently poor are likely to remain poor, suggesting a strong type of poverty trap that is qualitatively different from a dynamic thresholds-type poverty trap. While all types of households face static asset holdings, higher castes, larger landholders and more educated households are significantly less likely to be poor.

1 Introduction

Alleviating poverty is one of the key challenges for the new millennium. Meeting this challenge requires effective poverty reduction policies. Designing these policies, in turn, requires an understanding of the underlying welfare dynamics that determine how individuals and households escape or fall into poverty over time. Policy makers need information on two key issues: First, what are the levels of well-being that households are expected to reach over time and does this level of well-being differ across types of households. One can think of these levels as dynamic household welfare equilibria that households can reach given current economic opportunities and returns to their assets. Second, how do households move towards these equilibria? Does their well-being improve or worsen steadily or are there potential non-linearities in their underlying welfare dynamics potentially with associated dynamic poverty traps?

Precisely identifying the level and shape of household welfare dynamics has very practical policy implications. If there is but one dynamic equilibrium, the key questions then would be what this equilibrium level of welfare is relative to the poverty line and how quickly households move towards it. If it is sufficiently high for households to escape poverty, then policy can focus on speeding up the convergence process. In contrast, a dynamic equilibrium below the poverty line would suggest that eventually all households are expected to be trapped in poverty. Overcoming such a structural poverty trap would require structural changes that provide new economic opportunities for households that raise their equilibrium level of welfare.

If, instead, there are multiple dynamic welfare equilibria, a household's long-term welfare depends on its initial condition. If it starts above a dynamic threshold, in expectation it will move towards a higher level of welfare. A starting position below the threshold would put it onto a path towards another, low-level equilibrium. If this lower level of welfare lies below the poverty line then the threshold point would constitute the entrance to a second type of poverty trap. Clearly, the policy response in such a world would differ markedly from the single equilibrium case. It would require social policies to lift households above the threshold point and social protection measures to ensure that households don't fall below the thresholds in the aftermath of adverse shocks. A short term public investment in these social policies could harness the dynamic welfare process and yield large long term welfare benefits. Again, for the policies to be efficient, we would need to identify the precise location of the threshold point.

To get at these issues this paper employs the novel semiparametric panel data estimator developed in Naschold (2009). It tries to identify the shape of welfare dynamics and precisely locate dynamic welfare equilibria for households in three villages in semi-arid peninsular India. Case studies in the small existing empirical literature on household level welfare dynamics have focused on Sub-Saharan Africa. This paper contributes the first case study for India using the newly expanded ICRISAT Village Level Studies panel dataset which now spans 27 years with 13 observations per household. The long time span and the frequent observations make these data ideally suited to exploring the shape of long-term household welfare dynamics and their associated equilibria.

The empirical results suggest that these village economies are characterized by economic stasis. Levels of household well-being are effectively static. Absent any shocks, over time household asset holdings follow a random walk where a household can expect to remain at its current level of asset welfare. Crucially, this holds throughout the welfare distribution: the poor stay poor, and the non-poor stay non-poor.[1]

The remainder of the paper is organized as follows. The next section summarizes three competing theories of household welfare dynamics and provides a stylized theoretical framework that guides the analysis in this paper. Section 3 reviews the small empirical literature on modeling non-linear household welfare dynamics. Sections 4 and 5 introduce the data and construct the asset index that is needed for the subsequent asset dynamics analysis. Section 6 provides a summary of the econometric methods. Results are presented in section 7. Section 8 concludes.

2 Theories of Welfare Dynamics

Three main hypotheses from the macroeconomic literature on growth dynamics can inform the analysis of micro-level dynamic poverty traps: unconditional convergence, conditional convergence and multiple dynamic equilibria (2006).

The concept of unconditional convergence originates from the Solow growth model. In the context of household level dynamics it suggests that all households eventually gravitate to the same long term equilibrium, based on the assumption that asset dynamics for all households follow a common, concave, monotone Markov process. The dynamics underlying the conditional convergence hypothesis are the same. It expands the unconditional convergence concept simply by allowing exogenous subgroups to have a different dynamic path and equilibrium.

A priori, there is no clear reason why asset dynamics should follow an autoregressive process of this form. On the contrary, at least four theoretical models suggest that different types of nonconvexities can result in multiple dynamic equilibria and poverty traps if the lower stable equilibrium is below the poverty line.

First, the efficiency wage hypothesis (Mirrlees 1975; Stiglitz 1976; Dasgupta and Ray 1986; Dasgupta 1997) links worker productivity and earnings. Only if a worker can afford to consume more than a minimum level will he be productive and, hence, employed. Others who are unable to afford the minimum level of consumption remain poor. Second, limited access to credit or formal and informal insurance can limit a

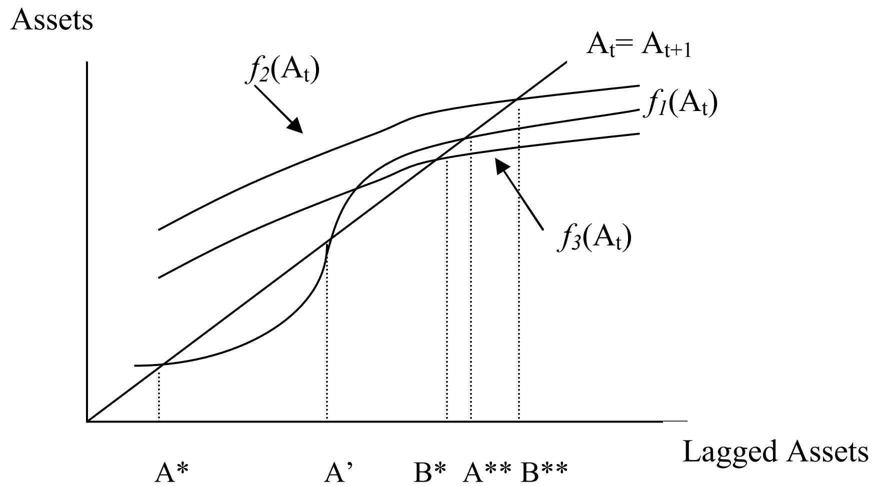
[1] In the context of these villages one cannot speak of 'rich' households, as even the richest households would be considered relatively poor in most contexts.

household's ability to invest in human capital (Loury 1981; Galor and Zeira 1993) or in an income-generating opportunity (Banerjee and Newman 1993). As a result any household dynasty starting below a certain level of wealth, or suffering a shock large enough to let it fall below this threshold, will be trapped in poverty. Third, if participating in society and finding employment require minimum levels of expenditure (Bradshaw 1993; Parker 1998), then poor households can be permanently 'socially excluded'. Fourth, child labor models (Basu 1999; Emerson and Souza 2003) suggest that poor households that have to send their children to work instead of school are trapped in intergenerational poverty since as adults these children do not possess the qualifications to access opportunities to escape poverty.

All these theoretical models have similar policy implications: if there are multiple dynamic equilibria with one stable equilibrium below the poverty line then the misfortune to start with low asset holdings or the realization of downside risk are structural causes of chronic poverty. Conversely, poverty traps and long term poverty could be eliminated if every household can be lifted above the unstable equilibrium threshold and if safety nets ensured that they remained there. Hence, one-off social expenditures would not only benefit households in the current period, but also result in higher welfare in all future periods. Current social expenditure would yield high long run returns.

The above theoretical models can be stylized by recursion diagrams in household asset space as shown in Figure 1. The recursion functions denote expected household asset accumulation paths. The horizontal and the vertical axes display household asset holdings in the previous and in the current time period, respectively $A_{t+1}$and A_t. Any point on the 45-degree line represents a dynamic asset equilibrium. Function $f_I(A_t)$ illustrates the case of multiple dynamic equilibria where the dynamic asset accumulation path crosses the 45-degree line several times. A precondition for the existence of multiple equilibria are non-convexities over at least a part of the asset domain. If the poverty line lies above A* then the unstable equilibrium point A' indicates a dynamic asset poverty threshold. Above this threshold point and absent any negative asset shocks households can be expected to accumulate further until they reach the high level long-run equilibrium point A**. Below A' households are on a trajectory which, in expectation, makes them poorer over time, moving towards the low-level poverty equilibrium at A*.

Figure 1 Stylized Asset Recursion Diagram for Different Shapes of the Asset Accumulation Path



Since these threshold points represent unstable equilibria a priori we might expect them to lie in a low density region of the distribution (Barrett and McPeak 2005), because in equilibrium households would be at either of the two stable equilibrium points. If this is the case we need estimation techniques that can identify these unstable equilibria using relatively few data points around A'. Fully parametric techniques would be at a clear disadvantage as the estimated polynomial function would be driven by the mass of observations around A* and A**, likely leading to imprecise estimation of threshold points such as A'.

The first alternative hypothesis of unconditional convergence can be represented by expected the asset recursion function $f_2(A_t)$. This would be consistent with a structural poverty trap if B** lies below the poverty line. The second alternative hypothesis, conditional asset convergence, would imply one such function for each exogenously determined subgroup. In the analysis below subgroup membership is defined by caste, landowning class, location of the household or its education level. Figure 1 illustrates the case of two subgroups. One follows $f_2(A_t)$ while the other is on trajectory $f_3(A_t)$ with each function having its own distinct dynamic accumulation path and asset equilibrium. As a result, in the conditional convergence case poverty traps could be subgroup specific.

Even in the absence of multiple equilibria and poverty traps, there may be a case for helping the poor escape poverty through redistributive policies that i) benefit them in the form of immediate transfers and ii) raise mean asset levels in subsequent periods. We can express future mean asset holdings as a function of households' current assets. If this function is strictly concave, as indicated by $f_2(A_t)$ and $f_3(A_t)$ in Figure 1, then future mean assets are a strictly quasi-concave function of households' current assets. Therefore, reducing current asset inequality would increase mean future asset levels (Aghion *et al.* 1999; Banerjee and Duflo 2003).

Again, this would imply that redistribution can support poverty reduction if the gains for the poor from redistribution are larger than any potential negative effects on economic growth. Testing for concavity of the recursion diagram using household data is therefore a micro-level test analogous to the test for the effects of inequality on economic growth in the cross-sectional macro literature (for a summary see Banerjee and Duflo (2003)).

Finally, it is worth noting that none of the above theories consider the speed of adjustment back to equilibrium as that is ultimately an empirical question. The models can only assume that households are 'temporarily' away from their respective stable dynamic equilibria (A* and A** in the case of multiple dynamic equilibria, B* and/or B** for unconditional and club convergence). This temporary deviation would be due to shocks causing asset losses or gains. Of course, in reality, depending on the speed of adjustment, this temporary state could be unacceptably long and justify policy intervention.

3 The Empirical Literature on Modeling Welfare Dynamics

Compared to the well-developed theoretical literature on welfare dynamics there is a relative dearth of empirical studies. The paucity of this literature is primarily due to the lack of suitable household panel data, but also to the empirical difficulties involved in modeling household welfare dynamics.

In terms of estimation methods the few existing studies have modeled household welfare dynamics either fully parametrically or nonparametrically. Existing parametric studies have limited themselves to a first order autoregression model. While longer lags could affect the dynamic welfare path, they also reduce the number of usable observations and use up degrees of freedom in the estimation.

Three published studies have used a model of this form. Two use the flow variables income and consumption to measure household welfare; the third is based on the stock variable of household asset holdings.[2] For Hungary and Russia, Lokshin and Ravallion (2004) estimate a third degree polynomial in income levels. Jalan and Ravallion (2004) use a fixed effect model in differences for rural China. Using income rather than asset data, neither of these two studies finds evidence for multiple dynamic equilibria. Both papers conclude that current income is a slightly concave function of lagged income. Therefore, poorer households would take longer to adjust to an income shock and are expected to move towards the single equilibrium more slowly than richer households. In contrast, Barrett *et al.* (2006) use asset data from Northern Kenya to estimate changes in assets as a fourth degree polynomial function of past assets, controlling for household and time specific effects. They detect nonlinear asset dynamics with one unstable threshold point and two stable equilibria suggesting the existence of dynamic poverty traps.

[2] Each measure has advantages and drawbacks for analyzing welfare dynamics. The beginning of section 5 explains why I chose household assets as the welfare measure in this paper.

One key problem with such parametric specifications is that if the unstable threshold points lie in an area with few observations, which the theories reviewed in the last section suggest, we need a large enough sample size that the fitted polynomial function can accurately reflect the few observations around the thresholds. If the sample size is too small the observations near the threshold point may not be picked up by the polynomial, but instead enter as heteroskedastic and positively autocorrelated error (Barrett 2005). Also, while high order polynomial functions present a way to adjust the coefficients so that in the centre of the domain the function exhibits the desired nonlinearities, they can make the function move around wildly towards in the tails of the distribution. This is to be expected from statistical theory (Hastie *et al.* 2001) and indeed is what Barrett *et al.* (2006) find in practice.

Three studies have tried to address these problems by using nonparametric estimation techniques. For Northern Kenya Barrett *et al.* (2006) run locally linear nonparametric bivariate LOWESS regressions of current herd size on its three month lagged value. Lybbert *et al.* (2004) run the same type of nonparametric regressions but on one and ten year lagged herd size in Southern Ethiopia. Adato *et al.* (2006) analyze household asset dynamics in South Africa using local regression methods.[3] All three studies using nonparametric techniques have found evidence for asset poverty traps.

Clearly, both estimation techniques used in the existing literature have limitations. Polynomial parametric techniques don't perform well with few observations around potential inflexion points. Nonparametric estimation is constrained in practice by how much it can control for other variables. Statistically, these two techniques mark the two extremes of the trade-off between the flexibility of the functional form and the ability to control for other covariates. Semiparametric techniques combine the advantages of parametric and nonparametric estimation and seem more suitable for modeling household welfare dynamics. This paper uses such a technique to analyze these dynamics for rural Indian households.

4 The Data

The data are taken from the International Crop Research Institute for the Semi-arid Tropics' (ICRISAT) Village Level Studies (VLS). The original first generation data (VLS1) was collected for the ten cropping years from 1975/76 to 1984/85. The cropping year runs from July to June. Here, I will refer to each year by the starting year only, that is, 1975 stands for the cropping year 1975/76. Collection of the second generation data (VLS2) started in July 2001 and is ongoing. The data released to date and used for analysis in this paper includes the year 2003.

[3] Their exact regression method is not specified.

The VLS1 data collection covered up to 10 villages in three states and a total of 400 households; the VLS2 spans 6 of those villages in two states containing some 265 households. The analysis in this paper is based on a subsample of these data selected on two main criteria. First, a household had to be included in both VLS1 and VLS2. This allows the construction of the longest possible panel spanning a period of 27 years. Second, income information had to be available for all years. This limits analysis to three villages: Aurepalle in the Mahbubnagar district of Andhra Pradesh, Shirapur in the Sholapur district of Maharashtra, and Kanzara in the Akola district of Maharashtra. This subsample contains 886 observations for 72 households with either 12 or 13 observations per household.[4]

Between VLS1 and VLS2 there has been some attrition as some households dissolved, while some others left the villages. Out of the 104 continuously surveyed household in VLS1, 72 could be included in the VLS2 sample. Ideally, we should try to control for the probability of attrition econometrically. This is only defensible if there is a variable in the survey that influences whether or not a household was resurveyed in VLS2 but which does not impact household income. However, the VLS was conceived primarily as an agricultural production survey. Hence, its module on household composition and characteristics is relatively small and does not contain any variables that can credibly identify the attrition probability (such as the place of birth, or the place of residence of relatives). The downside of not being able to control for attrition is that the results may not be representative for the villages. The attrition bias is likely to come from either end of the distribution: better-off and more educated households are more likely to migrate, while poorer households are more likely to disband, die off entirely, or merge with other households. A common factor in explaining attrition in other panels is the age of the household head. However, there is no reason to believe that the age factor affects household attrition differently for different wealth levels. Moreover, as Alderman *et al.* (2001) and Falaris (2003) show, even high levels of attrition in developing country panel surveys often do not affect the consistency of estimation. The upside of using only the continuously observed households is that the subsample containing these 72 households is a balanced panel. Thus, there is no further attrition bias during the period of analysis.

The VLS survey contains detailed information on key assets such as land, and agricultural and financial assets. Information on household composition and education is less detailed, but available at the basic level. The key feature that makes the ICRISAT VLS data suitable for exploring is the length of the panel. This is useful in two ways. First, the long time-span covered by the panels makes it suitable to track changes in assets which, absent any short term shocks, tend to be slow and may not be detectable in shorter panels. Second, unlike the few other panel datasets which cover similarly long periods[5] the VLS data have up to 13 observations per household. With that number it begins to be possible to estimate household specific asset dynamic curves using fixed effects.

[4] Some households are not included in all VLS2 rounds.
[5] For example, the Chilean data set by Scott with one observation in 1968 and another in 1986.

There is a 17-year gap between the last year of data VLS1 (1984) and the start of VLS2 (2001). There are different ways this can be handled. I annualized the change over this 17-year period to create a quasi-one-year period and use this in the asset dynamics analysis below.[6]

As a measure of material well-being I use household income rather than consumption. Although the permanent income hypothesis suggests that in theory consumption is a preferable measure of the economic standard of living, two key factors favor the use of income in the case of the VLS data. First, unlike consumption, income data are available for all years and all households. And second, there are reliability concerns regarding consumption data in the early years of data collection (Walker and Ryan 1990).

The VLS data collection was originally stratified into four equal sized landholding classes: landless laborer households and small, medium and large landowning households. The exact cut off points in acres between the landholding classes differ slightly between villages but are around 2.5 acres for small and 5.5 acres for medium landowners.[7] This information on landholding classes is used both in the regression analysis as well as to analyze asset dynamics by subgroup.

Caste membership is an important determinant both for the initial level of household well-being as well as for opportunities for economic advancement. For information on household caste membership I use Ryan's caste rank index (Walker and Ryan 1990) which classifies all castes into one of four groups with caste rank 1 containing the highest castes. Having four caste ranks reduces the number of necessary dummies in the subsequent regressions and makes the analysis of asset dynamics by subgroups more tractable.

Other variables are more standard and include household size, age of and years of education completed by the household head, and the number of working age adults and the number of children in the household. Variables are adjusted for household size and expressed in per adult equivalent terms[8] whenever appropriate. The construction of the asset index and the choice of variables used to construct this index are described in the next section.

---

[6] As a robustness check for this crude annualization I repeated the analysis dropping the 1984-2001 period altogether. The results did not change significantly. Thus, everything reported below uses the annualized data for 1984-2001.

[7] For more information on variable definition and on sampling procedures see Singh *et al.* (1985) and Walker and Ryan (1990)

[8] The adult equivalence scale was taken from (Ryan *et al.* 1984) who count men as 1 adult equivalent, women as 0.9 and children under 12 as 0.39 adult equivalents. In this paper I follow the VLS data standard and define children as anyone below 14. As a result the adult equivalent conversions using the Ryan *et al.* is incorrect for anyone aged 12 and 13.

5 Constructing the Asset Index

I use a definition of household welfare and poverty based on the stock variable asset holdings rather than a flow variable such as income or consumption. This focus on assets is for three reasons. First, the economic well-being of a household is dependent on its stock of assets. From a dynamic perspective it is the accumulation of assets which over time enables households to earn enough income to move out of poverty.[9] This makes an asset-based measure of household welfare more suitable for forward-looking policy design. Second, asset levels fluctuate less from day to day than income and, thus, are closer to the measure of structural well-being that is ultimately of interest to forward looking policy design. Assets can be interpreted as measuring the underlying structural well-being of a household whereas income, and to a lesser extent consumption, contains a much larger amount of stochastic variation (Carter and May 2001). Third, surveys tend to measure asset holdings more accurately than income or consumption. It is easier for a household to recall, and for enumerators to verify, how much X it has than how much it spent on Y or received in payment over the last fourteen days.

Unlike income or consumption, assets are multidimensional. Moreover, unlike income and consumption which is measured in monetary units and can, thus, be easily aggregated, assets are typically measured in different units, such as acres of land owned, years of education, or numbers of various types of livestock owned. To quantitatively analyze household asset dynamics we, therefore, need to summarize the different classes of assets into a single household asset index. This dimension reduction is particularly helpful when using nonparametric and semiparametric estimation techniques as it circumvents the curse of dimensionality and, thus, makes estimation possible with the small sample sizes commonly found in household survey data.

The asset index is constructed through a livelihood regression (Adato *et al.* 2006) which expresses household well-being as a function of household characteristics and asset holdings.[10] The fitted values of this regression can be interpreted as a asset index in which assets are weighted according to their marginal contribution to household *i*'s well-being.

Let household *i*'s subsistence need be the product of household size in adult equivalent units and the poverty line, and denote it by S_i. Further, let ℓ_{it} be a measure of *i*'s livelihood at time *t*, expressed as the ratio of its real income y_{it} to its subsistence need: $\ell_{it} = y_{it} / S_{it}$. Hence, ℓ_{it} measures a household's well-being in poverty line units (PLUs). This provides an intuitive normalization for ℓ_{it} and, hence, the asset index: A measure of 1 means that a household survives on an adult equivalent income right at the poverty line;

[9] Productivity and technologies and relative terms of trade are equally critical for a household's escape from poverty, but technologies and terms of trade are not state variables of the same sort as assets.

[10] An alternative way to construct an asset index is to use the first principal component from a principal components analysis of all relevant assets. Using the Indian VLS data and in earlier work on Ethiopia and Pakistan (Naschold (2006)) I find that the pattern of asset dynamics does not substantively vary across the two alternative methods of constructing the asset index. In the following I use the livelihoods regression methodology as the unit of the resulting asset index has a more intuitive interpretation (see next paragraph).

while $\ell_{it} < 1$ and $\ell_{it} > 1$ indicate poor and non-poor households, respectively. Household i's livelihood in PLUs can be expressed as

$$\ell_{it} = \alpha + \sum_{j=1}^{J} \beta_j \left(\mathbf{A}_{it}, \mathbf{C}_{it} \right) A_{itj} + \sum_{s=1}^{S} \gamma_s \left(\mathbf{A}_{it}, \mathbf{C}_{it} \right) C_{its} + \sum_{t=2}^{T} \delta_t T_t + U_i + \varepsilon_{it} \qquad \textbf{(1)}$$

where $U_i \sim_{iid} N\left(0, \sigma_u^2\right)$. A_{itj} and C_{its} are the j^{th} of J assets and the s^{th} of S household characteristics that contribute to i's livelihood at t. $\mathbf{A}_{it}$ and $\mathbf{C}_{it}$ are control vectors representing household assets and characteristics, and T_t are T-1 year dummy variables. Coefficients $\beta_j \left(\mathbf{A}_{it}, \mathbf{C}_{it} \right)$ represent asset j's marginal contribution to the livelihood. The parentheses indicate that the coefficients depend on all of household i's assets and characteristics at time t.

The fitted values yield the estimated household asset indices and are given by

$$\hat{\ell}_{it} = \hat{\alpha} + \sum_{j=1}^{J} \hat{\beta}_j \left(\mathbf{A}_{it}, \mathbf{C}_{it} \right) A_{itj} + \sum_{s=1}^{S} \hat{\gamma}_s \left(\mathbf{A}_{it}, \mathbf{C}_{it} \right) C_{its} + \sum_{t=2}^{T} \hat{\delta}_t T_t + U_i \qquad \textbf{(2)}$$

To estimate equation 1 we need to construct ℓ_{it}, which in turn requires setting a poverty line. The 1993 Expert Group of the Government of India suggested 49 Rupees per month in 1973/74 prices, which is around Rs630 per year per capita in 1975/76 prices. However, this would classify close to 90 percent of villagers as poor, including many households that would not be regarded as poor in the villages. Instead, I follow recent work by Badiani *et al.* (2007) and use a lower cut off point of 500 Rupees in 1975/76 prices as the poverty line.

Equation 1 was estimated using a random effects panel model using a second-order polynomial expansion of all assets and household characteristics. A random effects model is preferred for two reasons: First, a Hausman test against the fixed effects model did not reject the null hypothesis that the coefficients are the same. The probability that the critical χ^2 value with 74 degrees of freedom is greater than the observed χ^2 of 23.41 is equal to one. Second, the fixed effects specification explained only 65% of the variation in ℓ_{it}, compared to 74% for the random effects model; and the primary objective for the livelihood regression is to project the household asset index as precisely as possible.

The flexible second-order expansion for equation 2 allows the marginal returns to vary both with their own levels as well as with the levels of the other included asset and control variables. The estimation of equation 1 is based on the subsample of 72 continuous households. The subsequent asset dynamics analysis is, therefore, not affected by attrition though it may suffer from some selection bias.[11]

[11] As a cross-check I also estimated equation 1 by OLS and using the whole sample of 102 continuous and non-continuous households. This yielded a total four different household asset index estimates. The asset

The choice of explanatory variables was informed by the categories of assets identified in the livelihoods literature (see, e.g., Moser and Felton (2007)) and spans physical, productive, financial, natural and human capital.[12] Specifically, the asset vector A_i includes net financial assets, the value of houses, residential plots, productive equipment and consumer durables, acreage of dry and irrigated land owned, the number of bullocks and other bovine livestock owned. The household characteristics vector C_i contains the age and education of the household head, the household size in adult equivalent units[13] and the number of working age adults to proxy the labor endowment of the household. With the exception of the age and education of the household head, all variables were expressed per adult equivalent. To center the curvature of the polynomial at the sample mean, all asset and household characteristics variables and their squares and interaction terms were demeaned. Further, all variables expressed in Rupees, including the poverty line, were converted to real 1975/76 Rupees using the state-specific Consumer Price Index for Agricultural Laborers from Indiastat.

The main objective of this regression is to derive a set of weights to reliably project expected household well-being given its assets holdings. The focus of this regression is, therefore, less on the actual estimated coefficients but on the overall explanatory power. The regression has a good fit and can explain about three quarters of the variation in the household well-being as shown by an R-squared of 0.74. These results provide a solid justification for using the fitted values from equation 2 as the household asset index.[14]

Figure 2 plots the asset index on the vertical axis against its one-year lagged value on the horizontal axis. The asset index ranges from 0 to 23 PLU, with only 12 observations above 12 PLU. The data points are scattered fairly closely to the 45 degree line. This suggests a low level of asset mobility and is a priori evidence against nonlinear asset dynamics and multiple dynamic equilibria. This simple plot is consistent with overall economic stasis in these three rural villages.

---

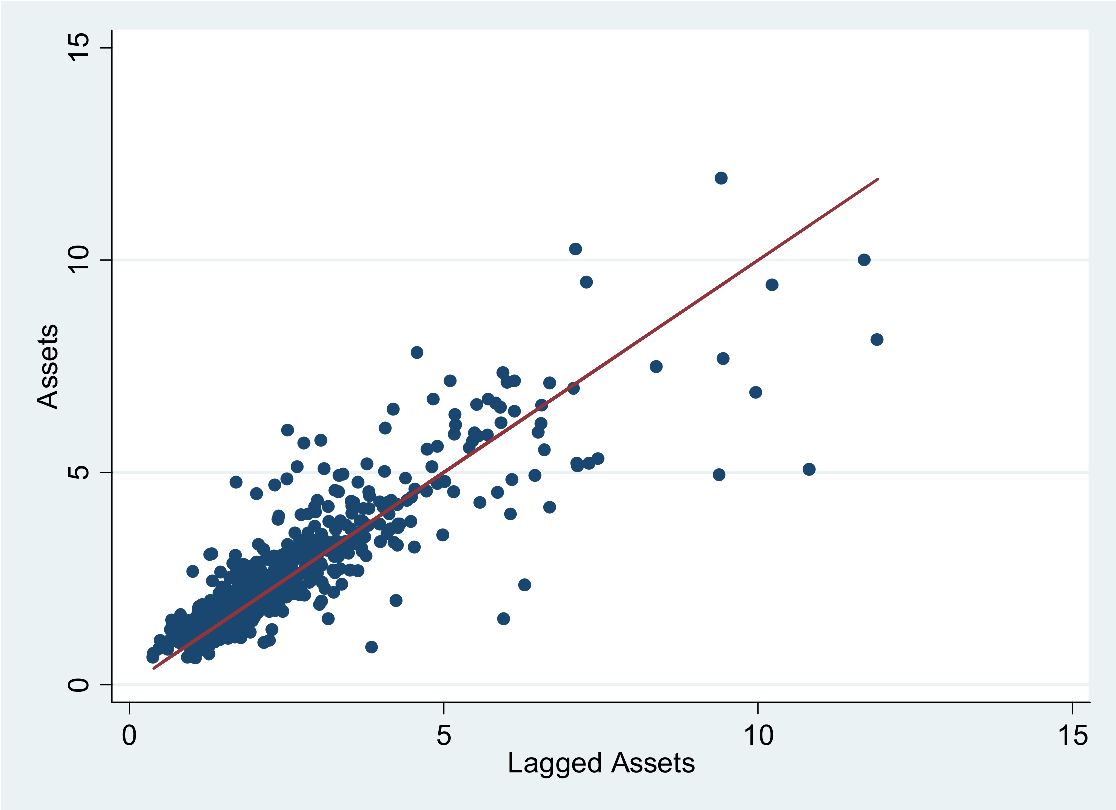
dynamics analysis using the three other asset indices (OLS all households, OLS continuous households and RE all households) produced substantively similar asset indices and asset dynamics results. The estimation results for all four specifications of equation 1 are given in Appendix A.

[12] The VLS does not provide sufficient detail on social capital.

[13] Strictly speaking I used the subsistence need which is equal to household size multiplied by the poverty line.

[14] Results of estimating equation 1 span about three pages. For brevity they are not included here, but are available on request from the author.

Figure 2 Scatterplot of Asset index against Lagged Asset Index (One year lag)



The constant from equation 1 is around 2, suggesting that the average level of household well-being for a household without assets is around twice the poverty line. The median household PLU for VLS1 (1975/6-1984/5) is around 1.5 and for VLS2 (2001/02-2003/4) is around 3.2. The relatively high average level of well-being is partially a result of four factors: First and foremost, household incomes far exceed consumption levels in the VLS. While there is no conclusive explanation for this, it is likely due to an underestimation of consumption levels (Townsend 1994; Morduch 2004). As a consequence, the magnitude of the regression constant of household PLU estimates is consistent with widespread consumption poverty headcounts of 76 percent for VLS1 and 22 percent for VLS2 (Badiani *et al.* 2007). Second, the period over which households have been observed is long. Thus, while average annual growth in assets is low at 0.8%, and consistent with economic stagnation, over almost thirty years this does represent significant compounded growth and has resulted in much higher income and asset holdings in VLS2. Third, due to the regression-to-the-mean effect, the fitted values of the asset index span a smaller range than the observed ℓ_{it}. This results in fewer projected poor households. Fourth, the Rs 500 poverty line is lower than in some existing papers.

6 Econometric Methods

In estimating household asset dynamics we face two main econometric challenges. First, we need to allow for any arbitrary form of non-linear asset accumulation patterns. This is best achieved by modeling the autoregressive asset relationship nonparametrically. Using a simple univariate nonparametric regression model we can fit a smooth function through a scatterplot of assets on lagged assets – as in figure 2 - without any assumptions on the functional form. Let A_{it} represent household i's asset index at time t then dynamic autoregression of household assets can be written as

$$\begin{aligned} A_{it} &= f(A_{it-1}) + \varepsilon_{it} \\ \varepsilon_{it} &\sim_{iid} N(0, \sigma_\varepsilon^2) \end{aligned} \tag{3}$$

Equation 3 can be estimated using a number of different nonparametric techniques. This paper presents results based on penalized spline regression. This method is borrowed from the statistics literature (Ruppert *et al.* 2003; Wand *et al.* 2005) and has not yet been used in applications in development economics. One obvious drawback of using simple bivariate nonparametric techniques is that we cannot control for other covariates.

The main analysis in this paper, therefore, employs the new semiparametric panel data estimator developed in Naschold (2009) again based on penalized spline regression. This estimator combines three desirable features for modeling household asset dynamics and is therefore conceptually superior to the techniques used in the previous literature. First, it retains the flexibility of the simple asset autoregression in equation 3 to flexibly model the asset autoregression to allow for possibly non-convexities in any part of the asset range. Second, it can control for other household characteristics, subgroups and time and location specific effects. And third, it can be extended for panel data estimation.

The full random household intercept mixed model representation of the this semiparametric penalized spline estimator takes the following form

$$\begin{aligned} \mathrm{A}_{it} &= \alpha + U_i + f\left(\mathrm{A}_{it-1}\right) + \mathbf{X}_{it}\beta + \sum_{t=3}^{T} \gamma_t T_t + \varepsilon_{it} \\ 1 \le i \le N,\ 2 \le t \le T,\ & U_i \sim_{iid} \mathrm{N}\left(0, \sigma_u^2\right), \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \end{aligned} \tag{4}$$

where $\boldsymbol{X}_{it}$ and all T_t enter the model linearly, and the relationship between assets and lagged assets is estimated nonparametrically. Vector $\boldsymbol{X}_{it}$ contains the following time-varying control variables: The age of the household head and its square control for life-cycle effects. Similarly, household size and its square control for economies of scale; and years of education of the household head and its square account for returns to human capital. The number of adults and its square and the number of children and its square control for available household labor and dependents, respectively. Also included in $\boldsymbol{X}_{it}$ are the two time-invariant household characteristics caste and landholding class. Continuous variables, their squares and interaction terms are demeaned to generate an exact second-order local approximation at the sample mean. Where appropriate, variables are used in per adult equivalent terms. T_t is a time dummy equal to one at time *t* and zero otherwise that accounts for time specific effects.[15] U_i is a random household intercept. Equation 4 is estimated using the *R* package SemiPar (Wand *et al.* 2005).[16]

7 Results

Table 1 reports the estimated locations of the stable dynamic equilibrium household asset holdings. The point estimates show only a single dynamic asset equilibrium for the rural Indian households and identify a very similar shape for the asset dynamics. This holds both for the nonparametric and the semiparametric penalized spline regressions and for the one year and the three lag structures.

The shape of the central tendency of households' asset dynamics is almost linear, suggesting that the VLS households accumulate assets very slowly, if at all. As illustrated by the scatter plot in figure 2 and as borne out by the recursion diagrams in figure 3 an figure 4 below, the VLS villages represent a stagnant economy. Households have a strong to tendency remain at their existing level of asset holdings and well-being.

The ranges reported in table 1 indicate the estimated dynamic asset accumulation path follows the 45 degree line very closely. In the first two columns the low and high values in brackets, respectively, show where the lower and upper 95 percent confidence bands cross the diagonal. The bracketed numbers, therefore, provide the 95 percent confidence interval for the asset equilibrium.

The confidence intervals for the nonparametric penalized splines regression span much of the asset range; from 3 to 9 poverty line units. In other words, we cannot reject the null hypothesis that, in expectation, no household's wealth changes from one year to the next. This represents a random walk along the 45 degree line and suggests a very strong case of economic stagnation. By controlling for covariates as well as allowing for flexible asset dynamics, the semiparametric penalized splines allow greater precision in the estimates. The confidence intervals for the dynamic asset equilibrium shrink to between 2.4 and 4.1 PLUs. However, this greater statistical significance does not detract from the economically significant result of very slow asset accumulation and overall economic stagnation for the VLS households, as even for the semiparametric techniques the asset estimated dynamic accumulation path still follows the 45 degree line very closely (see figure 4).

[15] The first time dummy is T_3 for the following reason. Since we are estimating a first order autoregression, the first observation for each household is at t=2. If we then omit the time dummy associated with this observation, i.e., T_2, to represent the base period, the first included time dummy is T_3. Similarly, the first included time dummy for the three year lag estimations is T_5.

[16] A summary of this semiparametric estimator is included in the Appendix. For more details see also Naschold (2009).

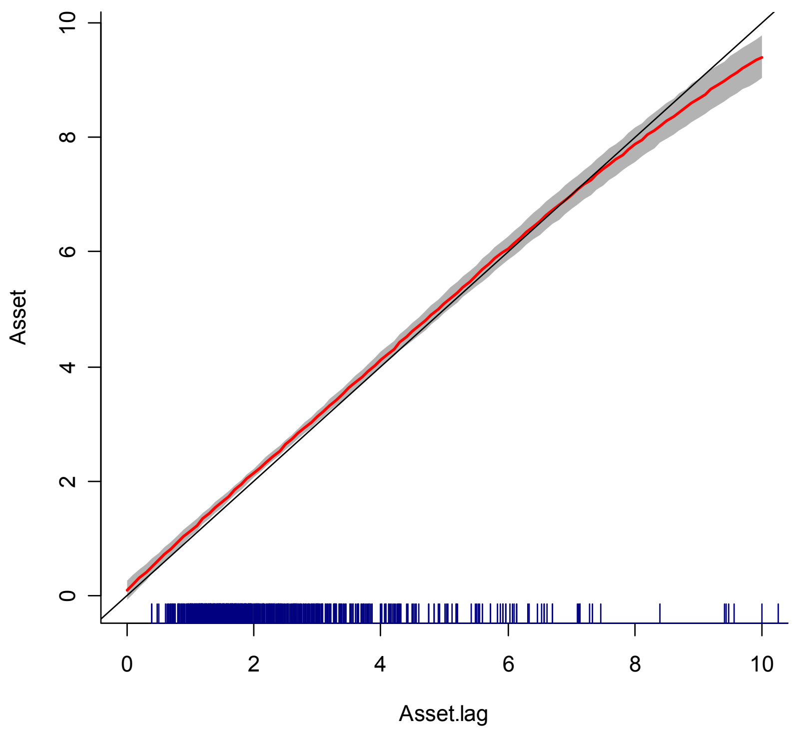
Table 1 Summary of Asset Equilibria Estimates by Estimation Technique and Lag Structure

	Approximate Location of Stable Equilibrium (in PLUs)		*% of observations in 95% Confidence Interval of the Stable Equilibrium*	
	1 year lags	3 year lags	1 year lags	3 year lags
Nonparametric Regression				
Penalized Spline (Equation 3)	7.1 [3,9]	5.7 [4.4,6.3]	21%	6%
Penalized Spline Random Effects	7 [3.6,9]	5.8 [5.4,6.2]	13%	2%
Semiparametric Regression				
PLM Penalized Spline	2.8 [2.5,3.6]	3.1 [2.4,4.1]	16%	21%
PLMM Penalized Spline RE (Equation 4)	2.8 [2.7,3.3]	3.2 [2.7,4.1]	9%	15%

- Asset Index estimated by Equation 1 using random effects.
- In each cell the line above the brackets indicates a single crossing point or its range.
- Ranges in brackets indicate where confidence bands overlap 45 degree line

Figure 3 shows the estimated recursion diagram for equation 3 using penalized splines. The central red line displays the estimated asset accumulation path. It is almost linear and follows the 45 degree line very closely. Hence, on average households do not move much from their initial asset position signifying an extremely low level of year-on-year asset mobility in the three villages. The grey areas above and below the curve show the 95 percent confidence bands. The rug plot at the bottom of the graph depicts the density of the observations.

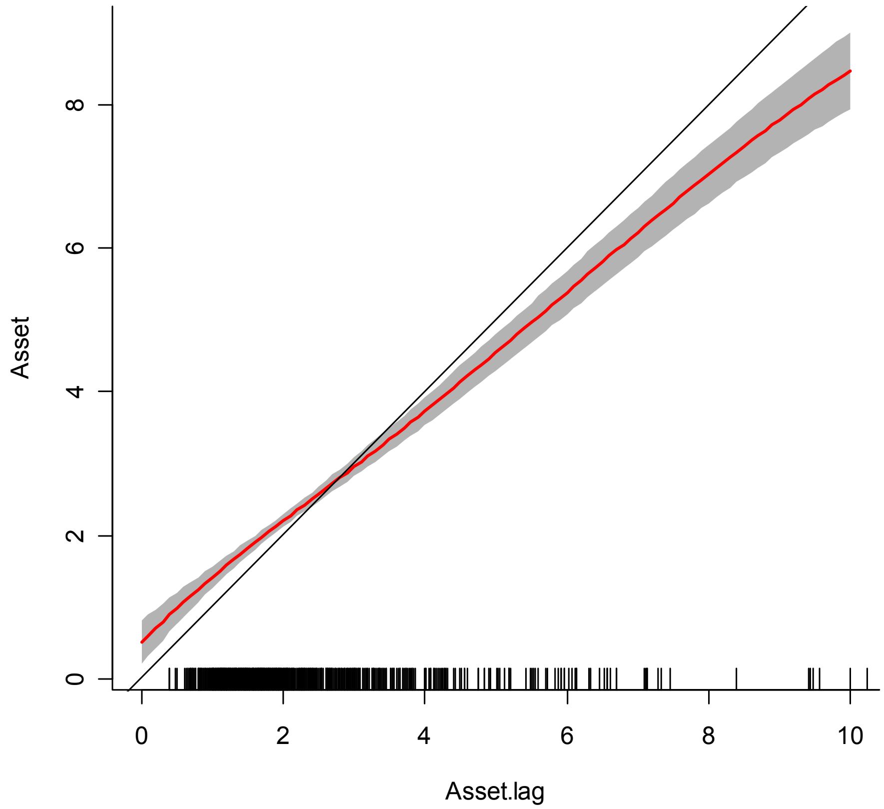
Figure 3 Assets vs. Lagged Assets (Nonparametric Penalized Splines) One year lag



The semiparametric penalized splines models produce asset recursions diagrams with the lowest dynamic equilibria of all the estimation techniques: between 2.8 PLUs for the one year lags and 3.2 PLUs for the three year lags. The semiparametric partially linear mixed model with the added random intercept from equation 4 is the preferred specification as it produces the tightest confidence bands of all the estimation techniques. Its asset recursion diagram is shown in Figure 4. To put these estimated equilibrium asset levels into perspective, the median household PLU level for 2001 was 3.1, suggesting that the median household had already reached its dynamic equilibrium at around three times the poverty line income, which translates into around Rs1500 per adult equivalent. For the reasons outlined at the end of section 5, this estimate is likely on the high side. For comparison, in 2001 22 percent of households still lived below the low consumption poverty line of Rs500 (Badiani *et al.* 2007).

Overall, the slopes from the semiparametric models are slightly smaller than for the nonparametric model, suggesting a somewhat faster, though still very slow process of asset accumulation.

Figure 4 Semiparametric Estimation of Assets vs. Asset Lagged (Penalized Splines Random Effects)
One year lag



In addition to the above analysis of the asset dynamics for the whole sample of households we can estimate asset dynamics and equilibria by subpopulation. Table 2 summarizes the results by village, caste rank, landholding class and education based on the semiparametric penalized splines estimation with random effects. All subgroups share a similar pattern for their estimated dynamic asset accumulation path. The shape of this pattern resembles that of the full sample discussed above and is characterized by a single dynamic asset equilibrium and a dynamic path that is relatively linear and close to the 45 degree line. The subgroups differ primarily in the location of their equilibria.

Among the three villages, Aurepalle has the highest dynamic asset equilibrium point estimate followed by Kanzara and Shirapur. The difference between Aurepalle and Shirapur is also statistically significant as their 95 percent confidence intervals do not overlap. This is consistent with the finding that income growth in the Andhra Pradesh village Aurepalle has outperformed the Maharashtra villages (Badiani *et al.* 2007) and suggests that economic growth in Aurepalle has managed to sustainably raise the level of structural well-being as indicated by a higher asset equilibrium.

Table 2 Stable Dynamic Asset Equilibria by Subgroups

	Number of observations	*Approximate Location of Stable Equilibrium (in poverty line units)*
		Semiparametric Penalized Splines (Random Effects)
By Village		
Aurepalle	286	4.4 [3,7]
Shirapur	338	2.5 [2.2,2.8]
Kanzara	312	3.5 [3,4.5]
By Caste Rank		
1 (highest)	260	4.2 [2.7,5.9]
2	195	3 [2.4,6.3]
3	247	2.5 [2.2,2.9]
4 (lowest)	234	2.1 [2,12]
By Landholding		
Landless	246	2.8 [2.5,3.5]
Small Landholders	265	2.5 [2.2,3.2]
Medium Landholders	206	3.3 [2.9,3.7]
Large Landholders	219	4.5 [4,7]
By Education		
No Education	544	2.2 [2,2.8]
Up to 4 years	204	4 [3,6]
More than 4 years	188	3 [2.5,5.5] and [7,15]*

- Asset Index estimated by equation 1 using random effects.
- * indicates that the equilibrium is driven by very few observations
- In each cell the line above the brackets indicates a single crossing point or its range.
- Ranges in brackets indicate where confidence bands overlap 45 degree line

The analysis by caste confirms the expectation that higher castes enjoy higher asset equilibria. The relationship between caste rank and the level of the stable equilibrium is monotone. However, the differences in dynamic equilibria are not statistically significant across caste ranks as evidenced by the overlapping confidence intervals in Table 2. This is due to a combination of heterogeneity within caste ranks and relatively small sample sizes for each caste rank.

Similarly, among landowning households, greater acreage systematically increases the level of the asset equilibrium. Large landowners enjoy statistically significantly higher asset holdings. The differences between the other groups are not statistically significant, again probably due to heterogeneity within landholding classes and due to small subsamples. Landless households have a slightly higher equilibrium than smallholders, but this is likely caused by shopkeepers and civil servants being included in the landless category.

Splitting the sample by education level of the household head we see that education raises the expected asset equilibrium level. The difference in dynamic equilibria between households with uneducated heads and those with up to 4 years of education is statistically significant. The highest education category has two equilibria. The one at around 3 PLU is likely driven by the larger number of household heads that have 5 or 6 years of education, thus, hardly more than the 'up to 4 years' group. The other, very high equilibrium is probably driven by the few household heads that have tertiary education.

I also tested the robustness of the above results against two alternative explanations of why we only see one dynamic asset equilibrium. First, social sharing rules may mean that any gains in assets by one household are at least partly distributed via its social networks. Alternatively, household composition may be endogenous. If a household manages to accumulate assets it also attracts people currently living outside of it to join the household. Both of these mechanisms would result in a household that is above the dynamic equilibrium to move back to it over time. We can test for these alternative explanations by re-estimating the livelihoods function in equation 1 but using total household income rather than household income adjusted by household subsistence needs as the dependent variable. This asset index then reflects any additional assets gained by the household, whether or not their returns were consumed by the (original) household members. Redoing the analysis with this asset index did not substantively change the asset recursion diagrams. This suggests that social sharing rules and endogenous household composition do not affect the asset accumulation path. This mirrors the results of similar analysis for rural Pakistan and Ethiopia (Naschold 2006).

A second reason why we might not see multiple equilibria is that the time period between observations is only one year. If total asset holdings change slowly, this may be too short an interval to pick up the long run asset dynamics. Indeed, the existing studies which have found multiple asset equilibria have either used longer spells [five year in South Africa (Adato *et al.* 2006) and thirteen years in Western Kenya (Barrett *et al.* 2006)], or are based only on pastoralists' livestock holdings, which are much more volatile than other asset holdings [see (Barrett *et al.* 2006) on Northern Kenya] or both [see (Lybbert *et al.* 2004) on Southern Ethiopia].

Rerunning the above analysis with a longer, three-year asset index lag did not substantively change the results. The asset recursions diagrams continue to show a single dynamic equilibrium at a level very similar to the one year lags. Again, this confirms the results for Pakistan and Ethiopia in Naschold (2006).

Finally there are two other possible, but untestable, explanations as to why the data do not show bifurcating welfare dynamics. First, evidence from Ethiopia (Santos and Barrett 2006) has shown that bifurcating equilibrium paths may depend on the quality of the growing season. When years are good all farmers expect to be on concave accumulation path. In contrast, in mediocre years only some farmers, including probably the experienced, expect to grow, whereas others expect to fall behind. In the VLS data we can partially control for good and bad growing years using information on the rainfall pattern. However, it appears that asset dynamics in the VLS data are not substantively different for good and bad harvest years, though this may be a result of rainfall being a crude measure for the quality of the growing season.

Second, the VLS covers poor rural populations. Close to 90 percent would fall under the poverty line recommended by the 1993 Expert Group of the Government of India. It is, therefore, possible that in the country as a whole there are additional higher asset equilibria, which are absent from the VLS data, as the VLS villages contain very few richer households. The only way to test for this would be to use a data set that is more representative of India as a whole. If indeed there were higher equilibria in other rural or urban parts of the country, then the findings in this paper could be interpreted as geographic poverty traps in the sense that there may be unique low level equilibria for the rural villages of the VLS.

8 Conclusions

This paper has examined household asset dynamics and tested for potential asset poverty thresholds in three villages in rural semi-arid India. The empirical results present a picture of economic stasis. Absent any major shocks households in these villages are expected to remain at their current level of well-being. This picture is confirmed by the low average annual growth rate in per capita asset holdings over the period 1975 to 2003 of 0.8%.

The existing theoretical literature on household welfare dynamics and most of the empirical case studies have focused on modeling the shape of the dynamic welfare path, in particular the existence of multiple dynamic equilibria and welfare thresholds. This is an important area of research as it can contribute to the design of more effective anti-poverty policies and, indeed, is what motivated the improved semiparametric estimation technique used in this paper. However, the empirical investigation of rural India showed clearly that in practice welfare dynamics can look very different from those hypothesized in the dynamic models and found in some of the existing empirical studies. The three rural Indian villages are characterized by a *lack* of welfare dynamics. Households' asset accumulation is a very slow, almost static process approximating a random walk along the 45 degree line. Since in expectation households remain at their initial asset positions, all households with an asset position below the poverty line are effectively trapped in poverty. The poor stay poor.

There is no evidence for non-convexities either for the whole sample or any of the subgroups. This suggests that households in the VLS villages do not face an asset poverty trap in the form of multiple dynamic asset equilibria. The estimated mean asset accumulation path also shows very little concavity. This has two implications. First, there is no significant difference in the speed of asset accumulation between asset-poor and asset-rich households. And second, reducing asset inequality would only provide direct benefits to the recipients, but would not enhance the overall rate of subsequent asset growth.

The mean overall asset equilibrium is around three times the lowest rural poverty line, or Rs1500 per adult equivalent per year in 1975/76 prices, or close to \$2 per day when converted to 1995 PPP \$. While this level is well above the poverty line and, therefore, may appear not too low, it is likely to understate the extent of poverty in the villages. Perhaps a better way of interpreting this result is in a relative context. The median household in the villages has already achieved the equilibrium level of well-being in 2001 suggesting that absent any changes in the welfare distribution. This suggests that the concurrent level of observed consumption poverty of around 22 percent is the long term equilibrium.

When we relax the assumption that households share a common underlying dynamic asset accumulation path and examine asset dynamics by subgroup we find predictable patterns of club convergence with a single equilibrium per subgroup. Higher castes, large landholders and more educated households have monotonically - and statistically significantly - greater asset equilibria than their lower caste, smaller landholder and less educated peers. In addition, recent improvements in economic conditions in Andhra Pradesh village of Aurepalle have resulted in a higher equilibrium level of asset holdings that in the Maharashtra villages.

Since the estimated asset accumulation path is linear in is central tendency and since even the best semiparametric model specification has confidence bands that straddle the 45 degree line of long term asset equilibria over much of the asset domain, there does not seem to be a way for households to move towards a higher equilibrium by exploiting non-linearities in asset accumulation. Instead, future improvements in welfare will have to come from structural changes that raise the asset equilibrium itself. In terms of social policy, the estimated linear asset dynamics also indicate that the key function for social safety nets is their traditional, most basic role of ensuring survival. Since there are no bifurcation points in the form of unstable dynamic equilibria safety nets cannot help leverage dynamic gains or, conversely, prevent dynamic losses.

References

Adato, M., M. R. Carter and J. May (2006). "Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data." Journal of Development Studies **42**(2): 226-247.

Aghion, P., E. Caroli and C. Garcia-Penalosa (1999). "Inequality and economic growth: The perspectives of the the New Growth Theories." Journal of Economic Literature **37**(4): 1615-1660.

Alderman, H., J. Behrman, H.-P. Kohler, J. Maluccio and S. Cotts Watkins (2001). "Attrition in Longitudinal Household Survey Data." Demographic Research **5**: 79-124.

Badiani, R., S. Dercon, P. Krishnan and K. P. C. Rao (2007). Changes in Living Standards in Villages in India 1975-2004: Revisiting the ICRISAT village level studies. CPRC Working Paper 85. Oxford, Department of International Development, University of Oxford.

Banerjee, A. and E. Duflo (2003). "Inequality and Growth: What can the data say?" Journal of Economic Growth **8**: 267-299.

Banerjee, A. and A. Newman (1993). "Occupational choice and the process of development." Journal of Political Economy **101**(2): 274-298.

Barrett, C. B. (2005). Rural Poverty Dynamics: Development Policy Implications. Reshaping Agriculture's Contributions to Society. D. Colman and N. Vink. Oxford, Blackwell.

Barrett, C. B., P. P. Marenya, J. G. McPeak, B. Minten, F. Murithi, W. Oluoch-Kosura, F. Place, J. C. Randrianarisoa, J. Rasambainarivo and J. Wangila (2006). "Welfare dynamics in rural Kenya and Madagascar." Journal of Development Studies **42**(2): 248-277.

Barrett, C. B. and J. G. McPeak (2005). Poverty traps and safety nets. Poverty, Inequality and Development: Essays in Honor of Erik Thorbecke. A. de Janvry and R. Kanbur. Amsterdam, Kluwer.

Basu, K. (1999). "Child Labor: Cause, Consequence and Cure with remarks on International Labor Standards." Journal of Economic Literature **37**(3): 1083-1119.

Bradshaw, J. (1993). Budget Standards for the UK. Aldershot, Avebury.

Carter, M. R. and C. B. Barrett (2006). "The economics of poverty traps and persistent poverty: An asset-based approach." Journal of Development Studies **42**(2): 178-199.

Carter, M. R. and J. May (2001). "One Kind of Freedom: Poverty Dynamics in Post-apartheid South Africa." World Development **29**(12): 1987-2006.

Dasgupta, P. (1997). "Nutritional status, the capacity for work, and poverty traps." Journal of Econometrics **77**(1): 5-37.

Dasgupta, P. and D. Ray (1986). "Inequality as a determinant of malnutrition and unemployment theory." Economic Journal **96**: 1011-1034.

Emerson, P. and A. Souza (2003). "Is there a child labor trap? Intergenerational persistence of child labor in Brazil." Economic Development and Cultural Change **51**(2): 375-398.

Falaris, E. M. (2003). "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Cote d'Ivoire, and Vietnam." Journal of Development Economics **70**(1): 133-157.
Galor, O. and J. Zeira (1993). "Income distribution and macroeconomics." Review of Economic Studies **60**(1): 35-62.
Hastie, T., R. Tibshirani and J. Friedman (2001). The elements of statistical learning: Data mining, inference, and prediction. New York, Springer Verlag.
Jalan, J. and M. Ravallion (2004). Household Income Dynamics in Rural China. Insurance against Poverty. S. Dercon. Oxford, Oxford University Press.
Lokshin, M. and M. Ravallion (2004). "Household Dynamics in Two Transition Economies." Studies in Nonlinear Dynamics & Econometrics **8**(3): Article 4.
Loury, G. C. (1981). "Intergenerational Transfers and the Distribution of Earnings." Econometrica **49**(4): 843-67.
Lybbert, T., C. B. Barrett, S. Desta and D. L. Coppock (2004). "Stochastic wealth dynamics and risk management among a poor population." Economic Journal **114**: 750-777.
Mirrlees, J. (1975). A pure theory of underdeveloped economies. Agriculture in Development Theory. L. Reynolds. New Haven, Yale University Press.
Morduch, J. (2004). Consumption Smoothing Across Space: Testing Theories of Risk-Sharing in the ICRISAT Study Region of South India. Insurance against Poverty. S. Dercon. Oxford, Oxford University Press.
Moser, C. and A. Felton (2007). The Construction of an Asset Index Measuring Asset Accumulation in Ecuador. Chronic Poverty Research Center Working Paper 87. Washington, DC, Brookings Institutions.
Naschold, F. (2006). Identifying Asset Poverty Thresholds - New Methods with an application to Pakistan and Ethiopia. Ithaca, Cornell University.
Naschold, F. (2009). Modeling household asset welfare dynamics: A semiparametric panel data estimator based on penalized splines. Ithaca, Department of Applied Economics and Management, Cornell University.
Parker, H. (1998). Low Cost but Acceptable: A minimum income standard for the UK: Families with young children: January 1998 prices. Bristol, The Policy Press and the Zacchaeus 2000 Trust.
Ruppert, D., M. P. Wand and R. J. Carroll (2003). Semiparametric Regression. Cambridge, Cambridge University Press.
Ryan, J. G., P. D. Bidinger, N. P. Rao and P. Pushpamma (1984). The determinants of individual diets and nutritional status in six villages of southern India. ICRISAT Research Bulletin No. 7. Patancheru, ICRISAT.
Santos, P. and C. B. Barrett (2006). Informal insurance in the presence of poverty traps: Evidence from southern Ethiopia. Ithaca, Department of Applied Economics and Management, Cornell University.
Scott, C. D. (2000). "Mixed Fortunes: A Study of Poverty Mobility among Small Farm Households in Chile, 1968-86." Journal of Development Studies **36**(6): 155-181.
Singh, R. P., H. P. Binswanger and N. S. Jodha (1985). Manual of instructions for economic investigators in ICRISAT's Village Level Studies. Patancheru, ICRISAT.

Stiglitz, J. E. (1976). "The efficiency wage hypothesis, surplus labor and the distribution of income in LDCs." Oxford Economic Papers **28**: 185-207.

Townsend, R. (1994). "Risk and Insurance in Village India." Econometrica **62**: 539-591.

Walker, T. S. and J. G. Ryan (1990). Village and household economies in India's semi-arid tropics. Baltimore and London, The Johns Hopkins University Press.

Wand, M. P., B. A. Coull, J. L. French, B. Ganguli, E. E. Kammann, J. Staudenmeyer and A. Zanobetti (2005). SemiPar 1.0. R package. http://cran.r-project.org.

Appendix A

Recall the extended partially linear mixed model from equation 4

$$
\begin{gathered}
\mathrm{A}_{it} = \alpha + U_i + f\left(\mathrm{A}_{it-1}\right) + \mathbf{X}_{it}\beta + \sum_{t=3}^{T} \gamma_t T_t + \varepsilon_{it} \\
1 \le i \le N,\ 2 \le t \le T,\ U_i \sim_{iid} \mathrm{N}\left(0, \sigma_u^2\right), \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)
\end{gathered}
\tag{4}
$$

where $\varepsilon_{it} \sim N(0,\sigma_\varepsilon^2)$, κ represents a knot, K is the total number of knots, T is the total number of time periods in the panel data, and the plus subscript $_+$ indicates that the term $\mathrm{A}_{it-1} - \kappa_k$ only enters the regression if $\mathrm{A}_{it-1} > \kappa_k$. Define

$$
\boldsymbol{\beta} = \left[\alpha, \beta\right]', \ \mathbf{X} = \begin{bmatrix}
1 & \mathrm{A}_{11} & X_{111} & \cdots & X_{11C} & T_3 & \cdots & T_{T-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \mathrm{A}_{1t} & X_{1t1} & \vdots & X_{1tC} & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \mathrm{A}_{n1} & X_{n11} & \vdots & X_{n1C} & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \mathrm{A}_{nt} & X_{nt1} & \cdots & X_{ntC} & T_3 & \cdots & T_{T-1}
\end{bmatrix},
$$

$$
\mathbf{Z}_{RE} = \begin{bmatrix}
0 & \cdots & 0 & (\mathrm{A}_{11} - \kappa_1)_+ & \cdots & (\mathrm{A}_{11} - \kappa_K)_+ \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & (\mathrm{A}_{1t} - \kappa_1)_+ & \cdots & (\mathrm{A}_{1t} - \kappa_K)_+ \\
1 & \cdots & 0 & (\mathrm{A}_{21} - \kappa_1)_+ & \cdots & (\mathrm{A}_{21} - \kappa_K)_+ \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
1 & \cdots & 0 & (\mathrm{A}_{2t} - \kappa_1)_+ & \cdots & (\mathrm{A}_{2t} - \kappa_K)_+ \\
\vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & (\mathrm{A}_{n1} - \kappa_1)_+ & \cdots & (\mathrm{A}_{n1} - \kappa_K)_+ \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & (\mathrm{A}_{nt} - \kappa_1)_+ & \cdots & (\mathrm{A}_{nt} - \kappa_K)_+
\end{bmatrix}, \textbf{ and}
\tag{5}
$$

$$
\mathbf{u} = \left[U_1, \ldots, U_K, u_1, \ldots, u_K\right]'
$$

The **u** vector contains two types of random effects: the $[u_1,\ldots,u_K]$ for the spline coefficients with $\mathrm{Cov}(\mathbf{u}) = \sigma_u^2 \mathbf{I}$ and the $[U_1,\ldots,U_K]$ for the random household intercept.

Then we can estimate equation 4 as the best linear unbiased estimator of the following mixed model representation of the penalized spline.

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z}_{RE}\mathbf{u} + \boldsymbol{\varepsilon}$$

$$\mathrm{Cov}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_U^2\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_u^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\varepsilon^2\mathbf{I} \end{bmatrix} \quad \textbf{(6)}$$

$\sigma_U^2\mathbf{I}$ measures the variation between households, $\sigma_\varepsilon^2\mathbf{I}$ measures the within household variation, and $\sigma_u^2\mathbf{I}$ controls the amount of smoothing used to estimate *f*. The partially linear model estimated by a mixed model representation of penalized splines combines the advantages of the global parametric model with the flexible functional form of a fully nonparametric model.

The random effects semiparametric penalized splines model in equations 4 and 6 was estimated using the *R* package SemiPar (Wand *et al.* 2005) which estimates the smoothing parameter λ through restricted maximum likelihood (REML), and cross checked using the model-selection-based algorithm in Ruppert et al. (2003: Appendix B).