Technology heterogeneity and poverty traps: A latent class approach to technology gap drivers of chronic poverty

Daniel Hill

University of Adelaide, Australia

Contributed paper prepared for presentation at the 64th AARES Annual Conference, Perth, Western Australia 12-14 February 2020.

Abstract

Empirical identification of the dynamics and drivers of chronic poverty, where households are unable to elevate themselves onto a positive wealth trajectory, have until recently provided limited insights when there is large heterogeneity amongst poor households. This paper redefines the use of a weighted asset index as a livelihood-generating technology using household panel data spanning 14 years in rural India. By recursively determining the livelihood technologies and grouping households based on their livelihood outcomes, I am able to show that there is significant heterogeneity in how households utilise asset holdings to generate a livelihood. This is complemented by exploring qualitative differences in the resulting groups, where livelihood technologies are associated with diversification of income sources and vulnerability to macroeconomic shocks. This indicates that stratification of groups through observed asset levels may misspecify the livelihood technologies households have access to. Finally, using the corresponding fitted values from the chosen livelihood technologies, varying shapes and convergence levels shows that accounting for heterogeneity is an important consideration when analysing poverty dynamics.

Contents

List of Figures

List of Tables

1 Introduction

In 2013, the World Bank Group (2016) estimated that 10.7% of the world population still lives under the standard poverty line of $1.90 USD a day, with an emerging consensus that this should be significantly reduced by 2030. Poverty intervention strategies have focussed on shifting households onto sustainable livelihood trajectories, defined as when a household has the capabilities to recover from stress and shocks (Scoones 1998). This has included large scale interventions, that combine productive asset transfers with training and safety nets to help build capacity and increase the likelihood this can be sustained (Banerjee et al. 2015). However, not every household in poverty is considered to be in chronic poverty (Barrett et al. 2006). Some may fall into poverty temporarily due to some shock but are able to return to a growth trajectory quickly. Given this, greater concern is on those in chronic poverty, where households suffer from persistent deprivations of assets and capabilities to overcome structural obstacles for an extended period of time (Moser 1998). Empirical analysis is valuable in understanding the levels of poverty, observed poverty dynamics, and the mechanisms behind why households have alternative livelihood dynamics. However, survey measures of income and consumption are often prone to large volatility and measurement error, making inference difficult in environments of large household heterogeneity.

To account for this, I focus on the livelihood regression index used in producing household livelihoods (Adato et al. 2006). This describes a relationship between a measured livelihood of a household, often income relative to a required poverty line, and the value of the assets owned. The estimation of this relationship gives the marginal contributions of assets in producing relative livelihood indices. A close resemblance can be found in production functions, where the set of marginal contributions is the technology available for households for assets to generate a livelihood. This is not an unreasonable assumption, as in the case where only productive assets are considered, then this form would directly represent a production function. However, it also considers that many other assets are important to the welfare of a household, such as an ability to save through exogenous shocks, maintain health, human capital and access alternative income generating opportunities. The fitted values from this estimated technology can be used to generate predicted livelihoods of households given a set of assets, avoiding the previously discussed volatility from using observed income or consumption.

In the case of Adato et al. (2006) and Naschold (2012), the livelihood index is estimated as an average across the entire sample population, and across all time periods, allowing for homogeneous differences in how livelihoods are generated between households and across time. This assumes all households share the same marginal elasticities from assets. The driving factors for the existence of chronic poverty must only be due to differences in the assets a household owns, or efficiency differences in the technology used. Although Adato et al. (2006) and Naschold (2012) address this question rigorously, consideration into how households use assets differently to generate a livelihood may indicate alternative results as to why some find themselves in chronic poverty.

This paper presents a latent grouping strategy that allocates households into the sub-technology that best represents their average observed livelihoods. This algorithm estimates the livelihood technology given an initial clustering on outcomes, and then reallocates households if the log-likelihood of being estimated in another group is higher. This is then repeated until there are no movements between groups in the sample. Similar methods have been used in the production literature, such as Greene (2005) and Orea & Kumbhakar (2004) in the banking industries, and Lin & Du (2013) in the Chinese energy efficiency. These papers all find variation in the technologies used within the industries in question, which were otherwise previously considered homogeneous in production technology or only differed on strict a priori grouping.

The resulting fitted values of the estimated livelihood indexes is used in a first-order auto-regressive process to consider the existence of possible unstable thresholds in livelihood dynamics as in Naschold (2012) and Adato et al. (2006) using further lags as controls to ensure that residuals are white noise. A unique panel data set covering 3 regions in rural India provides a reasonable cross section and length to ensure a large enough sample size for parametric asset dynamics, with 208 agrarian households observed across 14 years between 2001-2014. The data also includes detailed measures and values of household and productive assets. A large divergence in observed incomes over time (Rao 2008), accompanied with reasonable variation in household characteristics, income sources and expenditures, suggests differences in how households generate a livelihood. This motivates the case of heterogeneous technologies in livelihood dynamics further.

This analysis contributes to the literature in three ways. Firstly, significant heterogeneity in livelihood technologies are observed in the data, which highlights possible misspecification of the livelihood indexes in the previous literature. This is complemented with a qualitative analysis of the differences between the technologies, with large differences in households derive a livelihood from their diversification of income sources, volatility to macroeconomic shocks and access to collective technology such as irrigation. With the corresponding fitted values, conditional convergence is still observed as in Naschold (2012). However, these convergence levels occur at various levels, and trajectory functions display a variety of shapes across the domain of livelihoods. This shows that latent technology identification does matter when analysing poverty through livelihood trajectories.

The next section summarises the theoretical foundations of livelihoods, the purpose of identifying livelihood trajectories, and how this has been achieved in the previous literature. Section 4 then outlines the conceptual model of identifying livelihood technologies through latent grouping. Sections 5 and 6 give an overview of the data used, and introduces the empirical strategies used to identify groups and the trajectories. Section 7 reports the results for the livelihood technologies, and then reports of the characteristics of the resulting groups. Section 8 reports the resulting livelihood dynamics and how they differ between groups, and finally section 8 concludes the paper.

2 Poverty Dynamics and Measurement

This section presents the theoretical and empirical approaches in poverty dynamics in the previous literature, and how this has influenced the use of livelihood regressions as a tool for welfare dynamics. Fundamentally, chronic poverty occurs when households are unable to overcome structural obstacles for an extended period of time because of persistent deprivations of assets and capabilities (Moser 1998). Chronic poverty can also be thought of those with a low livelihood, where households lack the required assets and capabilities to recover from stresses and shocks (Scoones 1998). Multiple theories exist as to what the sources of these obstacles to obtaining a greater livelihood, which include labour market opportunities (Dasgupta & Ray 1986), nutrition traps in labour markets (Banerjee & Newman 1993), and human capital (Galor & Zeira 1993). The general conclusion of these theories show that there is not a pattern of unconditional convergence of household wealth over time, but structural features of the economy can drive those with lower endowments into chronic poverty.

Recently, Barrett & Carter (2013) argue that fixed costs and frictional credit markets are a possible driver for poverty traps in agrarian communities. Households with adequate wealth and incomes are able to afford more productive technologies, and can then maintain higher trajectories of livelihoods thereafter. Households without this wealth or income cannot afford the fixed costs of investment, and are unable to access credit to invest through debt. Livelihood dynamics under this poverty

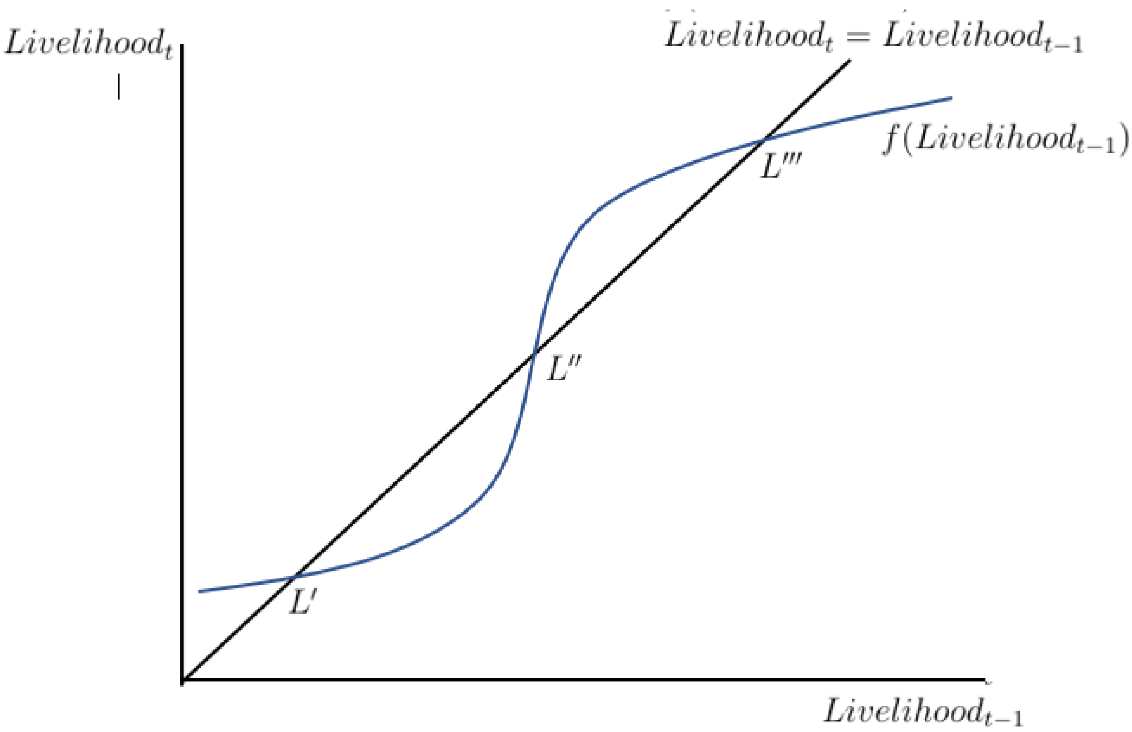

Figure 1: Livelihood Dynamics as described by Barrett et al. (2006)

trap indicate unstable thresholds in which households bifurcate towards either chronic poverty or an equilibrium above a given poverty line. This can be represented in figure (1), where any intersection between the 45°line and the auto-regressive function $f(Livelihood_{t-1})$ indicates a dynamic livelihood equilibrium. The equilibrium at L'' is the unstable equilibrium where households trend either towards a low livelihood profile of L', or climb out of chronic poverty towards L'''.

To understand whether households follow these livelihood dynamics, a comparative livelihood measure needs to be constructed. Income is a direct option of measuring the flows of welfare for a household. However, capturing average incomes in an annual survey is prone to measurement error, due to exogenous fluctuations such as seasonality and labour opportunities. This volatility also makes respondent recall and survey mis-measurement more likely, and difficult to generalise across the year.

Theoretically, consumption is an ideal alternative as it directly measures the capacity at which the household is sustaining their livelihood. The permanent income hypothesis then argues expenditures are smoothed over time, making extrapolation of consumption observations consistent. This can allow for measurement even in volatile economies such as agrarian communities (Deaton & Zaidi 2002). However, it is unlikely that those in poverty smooth consumption as expected during transient shocks because many poor households are faced with liquidity constraints and volatile expectations of future incomes (Chaudhuri & Ravallion 1994). Consumption in this case is vulnerable to measurement error, especially in poor agrarian settings where the value of food production that is consumed and not sold is difficult to aggregate. This is only compounded again by recall error and survey mis-measurement of expenditures that varies both in quantity and quality.

Indirect measures of welfare using assets have practical advantages. Assets reflect the underlying structural wellbeing of a household, as it measures both the household's accumulated wealth and the household's productive ability (Carter & May 2001). Although the permanent income hypothesis argues that wealth holdings should fluctuate given shocks to lifetime wealth, the previously discussed liquidity constraints amongst poorer households means that asset holdings are slow changing and less vulnerable to exogenous shocks (eg Carter & Lybbert (2012)). Asset based measures therefore can reliably determine the difference between transiently low income and persistent poverty. They are also easier to recall and count, which reduces seasonality and survey mis-measurement biases.

Given that the range of assets owned by a household are multidimensional, methods of dimension reduction are needed to ensure asset holdings are comparable across households. Filmer & Scott

(2012) compare three alternative approaches, a livelihood-weighted asset index, principal components analysis (PCA) and factor analysis. They find divergence in their asset indices in the case when it is benchmarked on household expenditures. This is supported by Michelson et al. (2013), where they find that although transitions and poverty rates are largely consistent across the indices, there was conflicting evidence of the existence of a poverty trap. As a result, care should be taken in the asset index chosen. Although PCA and factor analysis adequately reduces the dimensions of assets into a comparable measure, the elasticity weights of the assets on livelihood do not relate to weights associated with contributions to livelihoods. In contrast, the livelihood-weighted asset index as used in Adato et al. (2006) and Naschold (2012) describes a relationship between a measured livelihood of a household, often income relative to a required poverty line, and the value of the assets owned. Typically, it is based on a weighted sum of assets where weights are obtained from a second order flexible function estimated with livelihoods as an output and assets as inputs. These assets may include productive physical assets such as land and livestock, durable assets, financial assets and intangible assets such as human capital, available labour in the household and social assets.

Given that technology is defined as the set of input-output combinations that can feasibly be realised (O'Donnell 2008), the livelihood regression can instead be interpreted as a livelihood-generating technology rather than a method to reduce the dimensions of asset ownership. The household's technology set describes all possible livelihood outcomes that can be achieved given a set of assets available for the household. This is represented by equation (1). Household assets at time t, x_{it}, enter a generic technology set $T(\cdot)$ which maps into a measured household livelihood l_{it} (often income or consumption relative to household requirements). Time dynamics in the technology may also be considered (t). The first-order trajectory function then uses the estimated indexes to observe the possible existence of thresholds (equation (2)).

$$l_{it} = T\big(x_{it},\, t\big) \tag{1}$$

$$\widehat{l}_{it} = g(\widehat{l}_{it-1}) \tag{2}$$

In Adato et al. (2006), the index is used to identify an unstable threshold in rural South Africa, indicating many households with inadequate assets decline into a lower poverty level. Conversely Naschold (2012), Giesbert & Schindler (2012) and Naschold (2013) use similar techniques and do not find evidence of the existence of poverty traps, showing a possible slow convergence of livelihoods. However, these specifications build the livelihood index assuming the same marginal contributions of assets for livelihood across households and across time. Household fixed effects, used in Giesbert & Schindler (2012) (or stricter assumptions of random effects such as Naschold (2012)), allow for radial expansions along homogeneous technology sets, which only consider the case where households share the same asset elasticities but scaled by an efficiency constant. Year fixed effects, used in Michelson et al. (2013), Kwak & Smith (2013) and Naschold (2012), only allows for a homogeneous expansion (or contraction) of the technological frontier over time. Finally, separate elasticity coefficients have been estimated for each year (eg Michelson et al. (2013)), which enables technology to develop over time, but has only been considered homogeneously across all households. These assumptions give little consideration on how technological differences and change may contribute to livelihood dynamics over time. Given that influential research such as Solow (1957), Romer (1990) and Färe et al. (1994) explicitly highlight the role of technology on incomes and growth, greater considerations of technology in livelihood index analysis is essential for policy formulation. Critically from a development context, the technology also helps determine the process as to why some households generate greater livelihoods given certain assets compared to others.

Households face a number of financial, resource, societal and environmental constraints that may limit them from accessing the full range of feasible livelihood outcomes. For example, the marginal contribution of land would be expected to be different for a farming family with irrigated land, compared to a family that has poor quality soil and who instead derive much of their income as day labourers. It is unlikely a homogeneous technlogy will be able to capture these differences. Critically for poverty trap estimations, many of the studies that have used a livelihood regression attempt to observe poverty traps as described in theory, such as those described by Barrett & Carter (2013) where lower endowed households are restricted onto inferior technology profiles through credit market constraints. It is therefore inconsistent to estimate fitted values through a homogeneous technology to to test theoretical livelihood dynamics that are expected to be determined by heterogeneous technologies. Unless the technology differences are perfectly homogeneous or homothetic in nature, the estimated livelihoods are unlikely to be representative of true livelihood outcomes.

Considerations in technology heterogeneity can be derived from recent work in production theory (O'Donnell et al. 2008, Amsler et al. 2017), where heterogeneity is considered necessary to understanding unrestricted technology sets. Here, firms are faced with alternative feasible input-output constraints given differences in the available capital, infrastructure and production environment. These constraints shape technology sets available to the firm, which may be heterogeneous across groups. Examples include grouping technology across countries and estimating dairy farmers' technology and efficiency change in both South America (Moreira & Bravo-Ureta 2010) and Europe (Latruffe et al. 2017), and grouping according to region for energy efficiency in China (Wang et al. 2013).

From a household perspective, the feasibility of a technology set can be disaggregated into group-specific technology sets and individual technical efficiency within this technology set (O'Donnell 2016). The former highlights how households may be in chronic poverty because of group-specific input qualities, or are faced by environmental constraints on what they can achieve with these inputs. The latter refers to the case where households are considered technically inefficient due to management skill or available human capital. If either are drivers of chronic poverty amongst households, being able to identify and distinguish between the cases is essential from a policy perspective. This can be achieved through group-specific estimation of technology sets $T_j(x_{it}, t)$, where the marginal rate of substitution can differ across groups $j \in \{1, 2, ..., J\}$ for any combination of asset inputs. This follows the considerations made frequently in the production frontier literature, where technical inefficiencies (household and group fixed effects), group-specific constraints (alternative elasticities on asset inputs) and technological change (time fixed effects) can all be considered within the estimated technology sets. The heterogeneity in the estimation provides greater flexibility in the livelihood generating processes, reducing the risk of misspecification through the fitted values used in livelihood analyses.

3 Conceptual Model of Latent Livelihood Technologies

Now that the background and limitations of the previous literature has been discussed, the next section outlines the conceptual model of how this paper addresses these concerns when using a livelihood regression, and most importantly how to introduce heterogeneity correctly in the livelihood technologies.

Despite the estimation of heterogeneous technology sets having practical advantages for livelihoods, technology sets are fundamentally unobserved. Households could be incorrectly allocated into technology sets that do not represent their true technology. In the production literature, cohorts that share similar geography are often considered to share the same technology set, given that the legislative and geographical constraints are often major contributing factors as to why some are unable to reach an unrestricted technology set (Moreira & Bravo-Ureta 2010). In the case for livelihoods, a number of alternatives exist. Using the existing grouping strategies of Naschold (2012), prior endowments of land or social status are the defining barriers as to what technologies are available for the household. These are not unreasonable assumptions. In the case of land-holding differences, the existence of different technologies would indicate a possible poverty trap as described by Carter & Barrett (2006), where credit market frictions prevent those with poor initial endowments to afford the fixed investment for an improved or alternative technology sets. However, these are based on a-priori assumptions of what characteristics determine technological constraints. Given that technology can differ on a range of factors, these qualitative assumptions are likely to create errors in allocating households to their true technology sets (Orea & Kumbhakar 2004). Additionally, a core interest in poverty dynamics is on determining the differences as to why households with similar initial characteristics can have divergent livelihood trajectories. By stratifying households by these initial conditions, the researcher is restricting insight into these livelihood dynamics by not allowing households to differ based on outcomes.

I instead return to the underlying question of whether technology differences from assets is a driving factor for livelihood outcomes. A latent grouping of households, similar to the specification of technology heterogeneity by Greene (2005), can assign households to cohorts that best represent their observed outcomes. As outcomes are dynamic, households that share similar livelihood trajectories are considered to have shared outcome profiles. The iterative estimation approach using latent classes initially allocates households into cohorts using some base criteria and estimates an initial technology for each of those cohorts. Following this each household is compared to each technology using the likelihood function. Households are then allocated into the cohort for which the representative technology is most likely to be their own. If households are reallocated after the iteration, the process continues until no reallocation occurs between iterations, indicating convergence to a local optimum. This method less likely to be misspecified as it both allows technologies to differ across the sample whilst avoiding possible distortive assumptions associated with initial stratification.

The presence of significant heterogeneity in the functional form of livelihoods is shown when models which perform better with more than one cohort under standard parsimonious model selection tools such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) are observed. Multiple groups and different elasticities across the groups under the chosen model selection would imply that livelihood generating technologies vary in shape, and the assumption of homogeneous technologies is inadequate for estimating livelihood trajectories. This can be further compared with the a-priori stratification of groups used by Naschold (2012) to determine whether assumptions made in the previous literature are adequate in predicting sub-technologies. Alternatively, little differences in technology elasticities across groups would indicate that group-specific constraints to household technologies are of little concern in the sample. In this case, homogeneous technologies is a reasonable assumption, and radial differences along the technology sets through fixed effects is adequate for addressing sample heterogeneity.

4 Data

The data utilised in this analysis is from the Village Dynamics Studies in South Asia (VDSA) program, from the International Crop research Institute for Semi-Arid Tropics. Most studies previously researching poverty dynamics in India have depended largely on cross-sectional National Sample Surveys (NSS) collected in five year intervals (Thorat et al. 2017). The data presented here has the advantage of being collected at regular yearly intervals from 2001 and 2014 [1]

Sample bias from this attrition may occur when the poorer households merge with other households, and wealthier households are more likely to migrate to other regions. Naschold (2012) shows that there is little evidence of systematic attrition at either end of the wealth distribution in the years before 2003, which may alleviate concerns of later years. This is especially true given that Alderman et al. (2001) argue that attrition in household surveys from developing countries often has little impact on consistency estimates. Households who reported very high livestock, but with an income less that 10,000 rupees, are omitted because they might be herders who are reporting the livestock they manage and not the livestock they own. Observations with no income are also omitted as outliers. After this cleaning, an unbalanced panel of 205 households are available across all 14 years in the second wave, for a total of 2900 observations.

Adequate social and natural capital measures, which would be expected to have a role in the livelihood generating process (Moser & Felton 2009), are unavailable in this particular use of the data. Time invariant relationships of social (eg caste membership) and natural (eg geography) capital

[1] A first wave ran from the cropping season of 1975/76 to 1984/85. The first wave was not used due to concerns about reliability in the income and consumption data (Walker & Ryan 1990) and due to sample size where only 71 households are available across both waves of data collection.

would captured in the fixed effects constant if random effects is rejected, which reduces possible omitted variable endogeneity from these factors in the estimations.

As noted in a recent review of literature by Mosse (2018) and in notable studies by Thorat & Neuman (2012) and Gang et al. (2008), caste in particular would have dynamic and complex relationships with asset accumulation and livelihood outcomes. A fixed effects specification would not capture these dynamics as it only allows for a constant shift through time, leading to possible bias in the estimation. One option is to follow Ryan's Caste Rank Index (Walker & Ryan 1990), by grouping castes into one of four groups as a control variable. However, this can create a pre-stratification bias where the groups do not represent the true underlying interactions of caste membership and livelihood outcomes. There is also the trade-off with power by including further controls in the technology estimation. Without a viable alternative, a fixed effects specification is chosen, but it is noted that more work should be undertaken to translate complex caste dynamics into regression analysis.

Table 1 gives pooled summary statistics for all chosen asset variables, income and household characteristics. These assets combine measures of alternative productive assets, as well as durable, financial and human capital. All variables, where appropriate, are expressed in per capita equivalence for the household. This follows weightings from Ryan et al. (1984), which assigns one adult equivalence for adults and 0.4 for children.

Table 1: Summary Statistics

Statistic	Mean	St. Dev.	Min	Max
Total Income	133,032	162,801	100	2,485,547
Irrigated Land (acres)	2.78	3.65	0	42
Dryland (acres)	2.95	4.38	0	35
Large Livestock	18,229	31,765	0	335,000
Small Livestock	4,920	27,116	0	580,000
Buildings	143,536	154,199	0	1,500,000
Durables	84,718	173,449	0	3,278,160
Savings	23,858	63,776	0	1,076,911
Farm Equipment	26,182	76,953	0	1,242,100
Loans	55,461	96,823	0	1,186,000
Stock	5,829	10,171	0	157,666
Education (years in household)	24	17.12	0	124
Household Size	5.115	2.285	1	23

Note: All variables are expressed in value in Indian Rupees, unless otherwise stated.
2900 observations represent 205 households in an unbalanced panel

5 Empirical Model and Specification

The following section outlines the approaches used to estimate both the livelihood technologies and the livelihood trajectories. This includes choosing the correct specification of technology under the assumptions of homogeneity, and then utilising this in the latent grouping estimates. Endogeneity concerns also need to be considered, given that interpretation of the technology results is used to understand characteristics of cohort membership. Finally, a summary of trajectory functions in the previous literature is followed by the estimation strategy of trajectories in this paper.

5.1 Homogeneous and Latent Group Livelihood Technology Estimations

Firstly, an appropriate functional form for livelihood technologies needs to be chosen. Although alternative functional forms may be one way in which households may differ in their livelihood generating technologies, using the same technology has the advantage of being able to compare technologies directly. As a result, a baseline estimation using homogeneous technologies, as used in Naschold (2012) and Adato et al. (2006), tests for the most appropriate functional form. A translog estimation allows for greater flexibility for the marginal returns of assets to vary with its own levels and the levels of other assets, but the number of parameters required reduces the power of the estimation, especially with smaller sample sizes during the latent grouping estimations. As a result, a Cobb Douglas functional form, as chosen by other latent technology estimations (Orea & Kumbhakar 2004), is also considered. These can be represented by equations (3) and (4) respectively.

$$ln(l_{it}) = \alpha_i + \sum_j \beta_j ln(A_{ijt}) + \sum_j \xi_j ln(A_{ijt})^2 + \sum_{j,k} \nu_{jk} ln(A_{ijt}) ln(A_{ikt}) + \sum_t \gamma_{jt} t + \epsilon_{it} \quad \text{(3: Translog)}$$

$$ln(l_{it}) = \alpha_i + \sum_j \beta_j ln(A_{ijt}) + \sum_t \gamma_t t + \epsilon_{it} \qquad \text{(4: Cobb Douglas)}$$

Where i represents the households, j and k represent the household assets and t represents the year of the survey. l_{it} is the observed household income in per capita equivalence. β_j is the vector of elasticity coefficients for each asset j, which is then adjusted for scale and interaction effects in the translog specification. α_i represents the individual fixed effect when a random effect cannot be justified. The log is approximated through an inverse hyperbolic transformation, as zero values in assets would be undefined using a natural log. Where appropriate, time dummies are used as the time fixed effect, and the household mean is subtracted for the household fixed effect.

Complementing this homogeneous estimation is the latent grouping and estimation of households. Using the functional form most appropriate from the homogeneous estimation, livelihood technology parameters are estimated according to an initial grouping. The standard normal likelihood of a household existing in any of the estimated groups in a particular year is calculated, which is then averaged across all years to give a mean likelihood of belonging to a particular group. These likelihoods are converted to probabilities, and households are allocated into the group with the highest probability of their average technology. The process is repeated, using the new allocation, until there is no reallocation of households after an estimation. The grouping indexes of households indicate which livelihood group estimation best matches their realised outcomes. One issue with this estimation is that an average specification is vulnerable to households having outliers in their realised outcomes in a particular year. This is partially remedied by removing outliers with large and unexplained fluctuations in their income in a particular year, which allows for a better approximation of households that realise an expected flow of livelihood given their asset profile. To ensure statistical properties are maintained throughout the latent grouping, a minimum of 10 households are required to be allocated in each group. This is achieved by taking households that are next best represented by the groups under the minimum grouping size, and reallocating them until the minimum is achieved.

To initialise the latent algorithm, an appropriate initial grouping needs to be identified. A poor choice of initial grouping may allocate households to a local maximum of the likelihood functions, and not the global maximum. As a result, a number of alternatives are chosen to ensure the initial allocations are robust. Naschold (2012) chooses landholdings, village and education as possible stratification variables to split the sample into subgroups. These variables are therefore used as initial clustering. Despite this, there is no clear evidence that these initial groupings best represent final technologies, and therefore may not be robust to final allocations. As a result, mean landholdings, education, realised incomes and the mean fitted values of the homogeneous estimation are used in the Mclust (Scrucca et al. 2016) clustering algorithm in R, along with averaged fitted values for three year blocks from a fixed effects regression across the entire sample. The final allocations of all these initial groupings are compared to ensure final latent allocations are robust.

An OLS estimation is used within the grouping loop. Although this may not be the correct specification for households, the groups account for much of the heterogeneity of households that a fixed effects term identifies. A fixed effects estimation also does not allow groups to differentiate according to radial expanisons in their overall efficiency. Inefficient use of assets is one important technological difference between households, and the latent grouping should be allowed to differentiate households according to this constant shift [2]. After the groups are determined, the estimates can be re-estimated through a fixed effects estimation if appropriate, ensuring that any possible endogeneity between household fixed effect and the asset levels are accounted for. The optimal number of groups can be determined through the AIC and BIC statistics of each group, which punished estimations based on an aggregation of the total number of parameters across all group estimations.

Another consideration is how technologies develop over time. Elasticities may change over time amongst a given cohort, indicating a shift in the importance of certain assets in generating a livelihood. Michelson et al. (2013) achieve this by estimating alternative elasticities in their livelihood regression for the two periods they observe data. The VDSA data raises issues with this specification, as a shorter cross section but a longer time interval makes it difficult to guarantee reasonable power in the latent grouping algorithm. It also requires a shortening of this time interval, which may reduce the asymptotic properties of the fixed effects estimations. One advantage of the latent grouping method is that it allows for each group estimation to have alternative time dummy coefficients γ_{tj}. This allows mean differences for each year to be different for each group, ensuring variation caused by time to be controlled for adequately.

5.2 Endogeneity Concerns

Endogeneity in livelihood regressions has not been previously considered in the literature, given that the indexes have only been used as a tool to smooth transient income fluctuations. If the endogeneity effect can be assumed to be similar for all households, then the bias on the elasticity coefficients should not influence the final fitted values greatly. However, as this paper redefines the livelihood

[2] As a specification test, both random and fixed effects are used in the latent grouping, with some allocations failing to converge with larger groups. This is evidence that there was too little variation to differentiate groups within the latent grouping

regression as a livelihood technology, where the coefficients estimated are to be interpreted as unbiased and consistent estimators of the true marginal contributions an asset has on a livelihood, this endogeneity is an issue.

Firstly, endogeneity can arise in technology estimations when inputs and outputs are chosen simultaneously. In this case, households buy and sell assets to achieve a chosen income (or livelihood). A common approach in production literature is to use input prices as instruments. This is partially captured in the reported livelihood regression, where the value of assets derived from estimated market prices are used to proxy for quality differences. Land is measured in size because of inadequate data on land quality, it can be assumed that it is chosen ex-ante to livelihood realisations, meaning they are exogenous to this simultaneity. On the other hand, liquid assets such as savings and loans may be chosen simultaneously, where realised savings is a result of income flows. A lag net financial assets may be used instead of the realised net financial assets. This would ensure there is no simultaneity with realised income, and is a reasonable assumption given that savings and debt are often pre-determined in structural consumption-savings models. However, this requires dropping a year in the estimation. As including net financial assets in this case did not change the results greatly, dropping these financial assets was preferred over eliminating a year of observations.

Another endogeneity concern comes from measurement error in assets. Although measurement error occurs when deriving the exact value of the households assets, it is preferred given that the purpose of the asset index is to avoid otherwise transient fluctuations of income, which is expected to have greater measurement error when estimating livelihood measures. Finally, omitted assets that may contribute to livelihood and are correlated to the observed assets would bias the results. As described previously, if household fixed effects is chosen as the appropriate specification, time invariant variables are controlled for in this case. Time variant omitted variables are not be controlled for, and so care needs to be taken in the interpretation of the technology elasticities. However, this possible endogeneity does not matter in the trajectory estimations, as the fitted values is constructed the same regardless of which asset the variation is loaded onto.

5.3 Livelihood Trajectories

To model the dynamics of asset accumulation, a number of alternatives exist. A polynomial estimation similar to Jalan & Ravallion (2004) can be utilised. However, it may suffer from few observations around unstable threshold, where the estimation may identify these thresholds as heteroskedastic and autocorrelated error (Barrett & Carter 2013). This is the motivation of semi-parametric (Naschold 2012) and non-parametric (Barrett et al. 2006) estimation of dynamic livelihood trajectories, which can be more flexible around local non-linearities (Barrett et al. 2006). As described previously however, the sample size both in a cross section and across time alleviates the concerns that these less flexible estimations are unable to identify possible bifurcating behaviour.

The parametric estimation strategy chosen can be represented with equation (4), where $\widehat{l}_{ijt}$ represents the fitted livelihood values derived from a household's respective group estimation j in time t, p is the polynomial order applied to predicted livelihood in the previous period, and k is the additional lags used as controls. Polynomial expansions of up to eight, with lag lengths up to four, are tested. The most appropriate AIC and BIC statistic are chosen for the model specification used for each group, and tests are utilised to ensure that the errors are not serially correlated.

$$\widehat{l}_{ijt} = \sum_{p=1}^{P} \beta_p (\widehat{l}_{ijt-1})^p + \sum_{s=1}^{s} \beta_l (\widehat{l}_{ijt-s}) + \epsilon_{ijt} \tag{4}$$

Trajectories can vary between groups both in the shape of the function and as level differences. Using parametric methods that remove these level differences such as an Arellano-Bond estimation as in Jalan & Ravallion (2004) means that the possible levels in which intersection points are not observed. To account for these level differences, group technologies are re-estimated without household fixed effects, and the time levels are re-introduced into the predicted values. The trajectory specification then does not control for these levels, ensuring that the scale at which these influence the intersection points can be observed directly.

6 Livelihood Technology Results

Here I present the results from the homogeneous and latent estimations of livelihood technologies. This includes choosing the optimal model for the homogeneous technologies, and then the optimal group numbers for the latent grouping. This is then followed by robustness checks for the latent grouping, to ensure that results are stable with different initial clustering and model specifications. Finally, the results from the optimal groups are interpreted, to determine characteristics that are associated with cohort membership.

6.1 Livelihood Technology Model Selection and Robustness Checks for Latent Grouping

Results using the entire sample of equations (2) and (3) can be found in the appendix in tables (7) to (10). Hausmann and Breusch-Pagan tests for both models reject random effects, indicating significant unexplained fixed effects in the model. This is expected, given that time invariant characteristics of households such as caste membership, education of the household head and geographical location would both be endogenous with the asset ownership of a household, and are not controlled for directly. Time fixed effects cannot be rejected either, indicating large macroeconomic variation in rural livelihoods such as rainfall and input price volatility. As a result, two-way fixed effects models are preferred for the estimations both for the homogeneous technology, as well as the latent technologies post-allocation.

The adjusted within-R^2 values for both the demeaned Cobb Douglas and Translog functional forms favour a Translog specification slightly, explaining 25.5% compared to 25.1% when penalised for the number of parameters. The AIC and BIC statistics for both models are similar, and a preferred model cannot be drawn directly from this result. However, the restricted sample sizes in the grouping iterations reduces the power of the Translog model. It is further observed that the majority of elasticities in the Translog function are negative, which is unreasonable for technologies expected to have a positive return for the household. Given these concerns, a Cobb Douglas functional form is preferred given goodness of fit and power is maintained with fewer parameters to estimate during the latent estimations.

The test statistics which determine the optimal number of groups for the latent grouping algorithm are shown in table (2) for two to eight groups. The AIC and BIC statistics indicate that the optimal specification occurs with seven cohorts, meaning that additional group of eight fails to explain further heterogeneity at the cost of degrees of freedom. However, seven and eight groups occasionally fail to converge depending on the initial grouping specification, and is especially true when using two-way fixed effects. This is evidence of over-specification, and so six groups is chosen as the optimal number, and the results are reported in table (3). Each group has higher explanatory power than both the OLS model and the homogeneous fixed effects model, with a range of adjusted within-R^2 values between 0.282 to 0.514. This is a first indication that the latent groupings identified sub-technologies within the sample, which are better explained than through a homogeneous technology with just level differences through fixed effects.

Table 2: Latent Grouping Test Statistics

Number of Groups	2	3	4	5	6	7	8
Parameters Estimated	46	69	115	115	138	161	184
Log Likelihood	-3662	-2856	-2681	-2586	-2516	-2443	-2431
AIC	6217	5868	5545	5402	5308	5207	5230
BIC	6457	6199	5964	5901	5876	5849	5934

Note: Parameters estimated are the sum of all parameters for each group specific technology. Log Likelihood calculated as the sum of log-likelihoods for each observed year.

Table 3: Latent Grouping Elasticity Estimates on Observed Income

		Latent Group Estimates					
	Entire Sample	(1)	(2)	(3)	(4)	(5)	(6)
Irrigated Land	0.134**	0.048	0.225	0.224	0.366***	0.098	0.161**
	(0.054)	(0.072)	(0.141)	(0.234)	(0.119)	(0.119)	(0.077)
Dryland	0.118**	0.096	−0.0004	0.046	0.140	0.030	−0.014
	(0.057)	(0.072)	(0.165)	(0.257)	(0.118)	(0.129)	(0.077)
Large Livestock	0.005	0.006	−0.002	0.021	−0.0002	−0.027*	−0.004
	(0.006)	(0.007)	(0.023)	(0.024)	(0.012)	(0.016)	(0.006)
Small Livestock	−0.004	−0.020***	0.034*	−0.002	−0.0004	0.005	−0.006
	(0.006)	(0.007)	(0.020)	(0.025)	(0.008)	(0.017)	(0.007)
Durables	0.103***	0.198***	0.136*	0.098	0.060*	0.033	0.048**
	(0.022)	(0.029)	(0.073)	(0.108)	(0.035)	(0.068)	(0.022)
Buildings	0.025	0.038*	−0.035	0.054	−0.065***	0.029	0.154***
	(0.016)	(0.022)	(0.038)	(0.068)	(0.025)	(0.044)	(0.035)
Farm Equipment	0.053***	0.060***	0.067	0.031	0.019	0.299***	0.028*
	(0.014)	(0.014)	(0.054)	(0.086)	(0.026)	(0.059)	(0.014)
Education	0.167***	0.143***	0.255**	0.250	0.437***	0.336**	0.044
	(0.045)	(0.050)	(0.115)	(0.241)	(0.105)	(0.137)	(0.048)
Constant	−0.104*	−0.307***	−0.055	0.743***	−0.301***	−0.025	−0.176***
	(0.061)	(0.080)	(0.223)	(0.268)	(0.103)	(0.165)	(0.066)
Observations	2,900	728	266	328	446	404	728
R^2	0.257	0.482	0.282	0.500	0.514	0.512	0.494
Adjusted R^2	0.251	0.467	0.220	0.466	0.490	0.486	0.479

Note: Entire sample is the two-way homogeneous estimation for Cobb Douglas technology. Group estimates are two-way fixed effects models as well. Household fixed effects are accounted for through household level demeaned data, and time fixed effects use time dummies. Time coefficients are reported in appendix in table (11). Within-R^2 reported for all models. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

The probabilities of household membership into each group range from 0.51 to 1, with a median of 0.99. These are relatively high, which indicates that households are well represented by the group they are classified into, and not the alternatives. Overall, the results suggest that the methods used identified and allocated households into distinct cohorts, with strong evidence of heterogeneous livelihood technologies. To ensure the results are robust the the initial allocations, groupings using landholdings, education and villages from 2001 are also used and the final results are compared to the clustering allocation. The allocation through villages is fairly similar, with 36% of households allocated the same as the clustering algorithm. Results from the initial allocations of land and education only share around 15% of households, an allocation similar to what is expected under random assignment of six groups. This indicates that the latent algorithm may be fragile for data sets with strong heterogeneity, and that the initial allocation is important to ensure the estimations find a global maximum likelihood. What is also interesting is that about 60% of the households are allocated into the same group from the initial clustering algorithm and the latent technology estimation. This indicates that the clustering algorithm may be a reasonable shortcut to quickly allocate households into distinct groups, given the correct choice of information to cluster on.

To compare with the grouping chosen by Naschold (2012), the final latent grouping is compared to groups only stratified through observed initial allocations of land-holdings, village and education levels. The correlations are low, with the initial grouping of four levels of land performing the best by sharing 24% of the final grouping cohorts. This indicates that initial criteria to pre-determine technology sets poorly identifies the cohorts that households are actually members of. Stratification choices based on initial criteria are likely to misspecify households into incorrect technology sets, and supports latent grouping as a methodology as it avoids these subjective errors in identifying sub-technologies.

As the average cohort membership is used to estimate the true cohort membership across the years, this does not allow movements of households into different groups across years. Some groups may be well represented for all years, indicating that households within this technology set are unlikely to move to other groups. Conversely, some groups may only be represented for a fraction of the periods for many of the households, and so these group cohorts may have more volatile technology sets. To test this, for each household belonging in a group, the probability of belonging in this group compared to others for each year is also calculated, with the histograms of these probabilities shown in the appendix. The mean of these probabilities range between 21% and 34%, indicating that households are likely to be represented by alternative technologies for some of the years in the panel. Again, this is evidence that using the first period as a stratification of sub groups is an unreasonable specification, and that the mean probability is a better allocation measure to assign households into technology sets. This is supported by the histograms for the probabilities (figure 6 in the appendix), which are close to normally distributed for most groups, and so are unlikely to be affected by significant outliers in realised outcomes in a particular year.

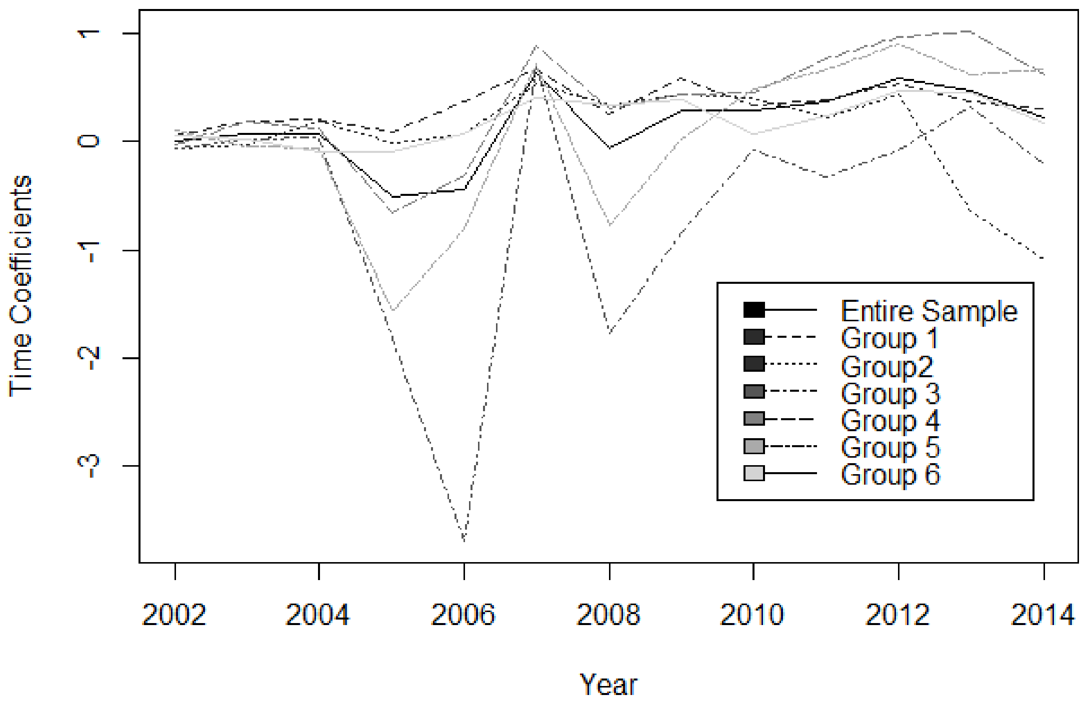

Figure 2: Time Dummy Coefficients by Group

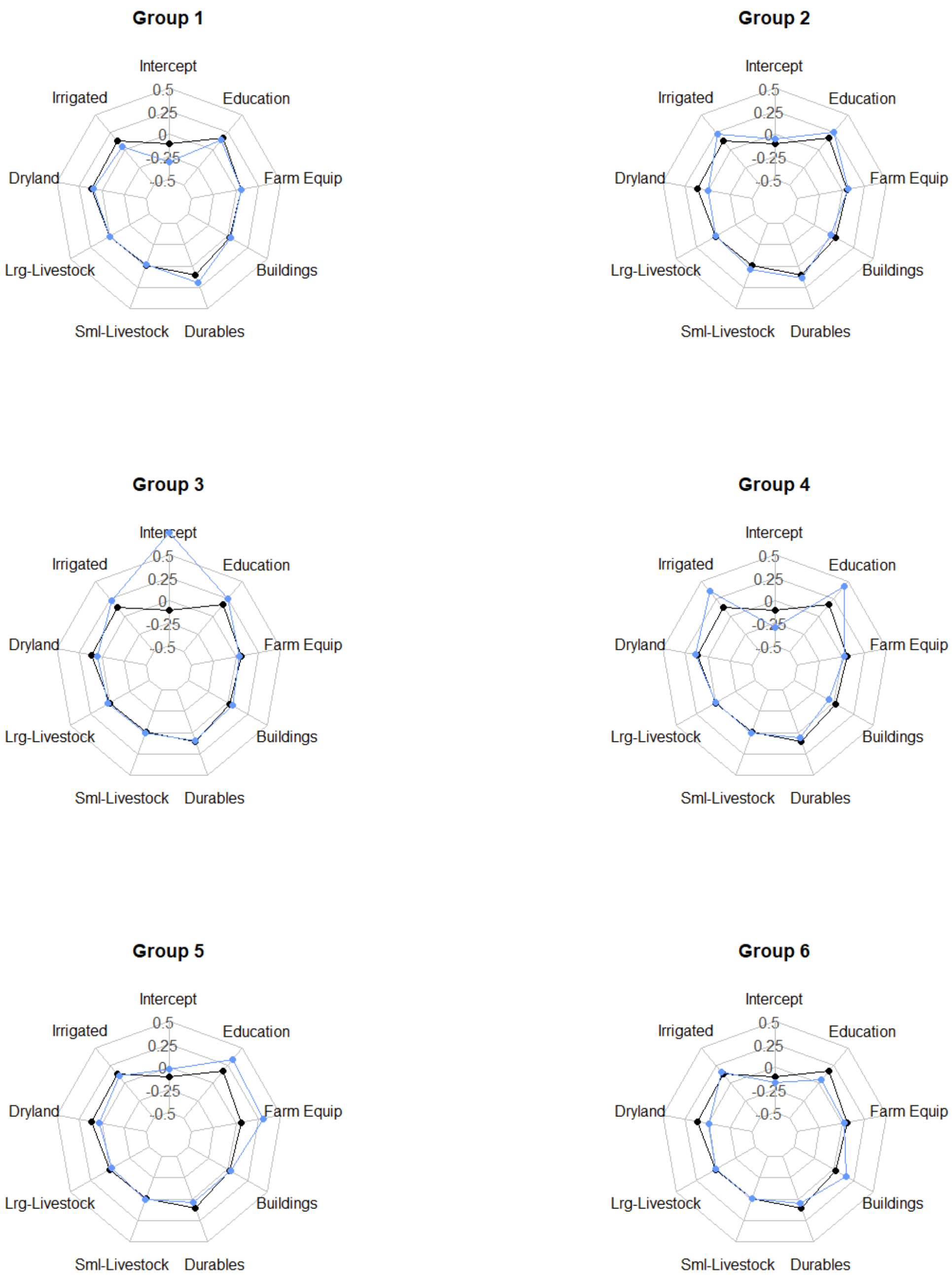

Figure 3: Radar Diagrams for Group Coefficients

6.2 Characteristics Associated to Livelihood Sub-Technologies

Now that the optimal latent grouping model has been determined and that it holds to the specification tests, this section interprets the resulting groups and technologies. Table (3) is accompanied by the radar diagrams in figure (3). These diagrams plot the elasticity coefficients for each group (the lighter blue technology) compared to the homogeneous estimation (the darker black technology). This allows to visually observe how technologies differ, and whether there are technology sets that dominate others in the sample. The estimated coefficients represent the output elasticity parameters in the Cobb Douglas function. For example, the 0.366 coefficient for irrigated land in Group Four means that for a 1% increase in irrigated land, on average and ceteris parabis, Group Four will have an increase in livelihood of 0.366%. As a result, different elasticities imply that households do not share a homogeneous technology set, and due to a variety of constraints households derive different levels of livelihoods when accumulating a particular asset. Technical livelihood inefficiencies, represented by the constant term, indicates how efficient a group is in their use in assets overall. The constant term represents all time invariant factors that influence livelihoods, such as geography and caste, and can be defined as the livelihood technical efficiency constant to align with production literature (O'Donnell 2016). Groups with different technical inefficiencies, but with similar elasticity coefficients, imply that these groups share a homogeneous technology set but one is more efficient in their use of assets overall compared to the other. As this term is the log efficiency constant, a positive constant occurs when a group has a technical efficiency constant greater than one, and negative implies a constant less than 1. Finally, the time dummy coefficients in table (11) in the appendix and plotted in figure (2) show how this technical efficiency may change year by year. Groups may share a homogeneous technology set, but may be more or less vulnerable year by year due to macroeconomic shocks affecting the overall productivity of assets. The following aims to understand in what way do the chosen groups differ along these three technological factors, and the possible constraints and sources that allocate certain households into these technology sets.

Group One has the highest median observed income out of all the groups estimated. Out of the statistically significant assets, they derive the greatest amount from durable assets out of any group, and derive a reasonable amount from changes in education in their household. Farm equipment is a small component of their livelihoods, and small livestock gives a minor negative return. They also have the lowest average constant term out of any group of -0.307, and have lower land stocks than groups three and five, and fewer large livestock. These results seem counter intuitive, as the wealthiest group would be expected to be able to derive more from their assets, and possibly have larger stocks of assets. However, they have the least volatile time dummies out of all the groups. This means that group one has a much more stable technology set compared to the other groups, where small homogeneous expansions and contractions in their technology sets indicates they are less vulnerable to exogenous macroeconomic shocks. They also have the largest income from non-farm sources out of any group, meaning that many derive more income from sources such as construction, skilled trade and government programs. This might mean that group one is represented by well diversified households, who derive incomes from a range of sources: a supported by findings from Deb et al. (2014), who concludes that much of recent development in the sample villages has come from strong livelihood diversification and growth in non-farm income.

Group Two have the least amount of household education out of all the groups, with the other groups having a median household education between five to eight years more. They also derive much of their income from farm labour and non-farm sources, but very little direct farm income. They are relatively efficient at converting their assets into an income, with a constant not significantly different to 1 and a reasonably stable time dummy trajectory up until recent years. This may be indicative of a number of households constrained in their opportunities, such as households in lower castes or with those with low productive land. These households may source their income mostly from caste-specific trades, accounting for the more stable income year by year.

Groups Three and Five are similar in many respects. Both groups share the largest land holdings both in irrigated and dryland, as well as the largest value of livestock and farm equipment. They also derive the most farm income out of all the groups, but receive very little income from other sources. This suggests that these two groups are represented by farming families. Where they differ however is in the constant parameter, which is made up of both the constant term and the time dummies year by year. Group Three has a highly volatile efficiency parameter, which falls by a significant amount in the years of 2005 and 2006, and again in 2008. The large and statistically significant constant means it may not be the least efficient out of all the groups in these periods, but shows that macroeconomic factors heavily influence realised incomes for Group Three. Group 5 follows the same patterns in the time dummies year by year, but at a smaller scale than Group Three. Given that Group Five derives greater returns from any changes in farm equipment and education than Group Three, it may be that Group Five use different technologies to alleviate changes in weather patterns or input prices, or use the land for different crops or livestock.

About half of Group Four is represented by households from the village of Kanzara, and a quarter from Shirapur, 300km Southwest. As Kanzara has had a large reduction in poverty due to improved collective irrigation facilities, social organisations and cultivation of high value crops (Deb et al. 2014), this helps explain the large and statistically significant coefficient on irrigated land. This is supported by the fact that a village nearby not represented well by Group Four, Kinkhed, has not benefited from irrigation due to a lack governance and social capital. Shirapur has also seen a reduction in poverty because of an expansion of cash crop cultivation. Improved market linkages in Kanzara and increased salaried work in Shirapur also help explain large returns from education.

Group Six has the lowest land ownership out of all the groups, and also the lowest amount of livestock. Despite this, they still earn a reasonable income, driven mostly by farm labour. This explains why they derive no statistically significant return from education, where they may face other constraints in social movement such as caste in utilising this education in other income-generating activities.

Overall, these groups show that the sample utilised is highly heterogeneous in their use of different assets to generate a livelihood, as well as efficiency differences on average and year by year. Groups Three and Five are the only groups that almost share homogeneous technology sets, and are mostly differentiated by volatility in overall technical efficiency. However, they still differ in their optimal use of assets, especially in farm equipment and education. Overall, this is clear evidence that the use of a homogeneous technology specification such as Naschold (2012) and Adato et al. (2006) misspecifies livelihood technology sets amongst households with large heterogeneity, even when accounting for radial expansions in technical efficiencies through fixed effects. As the groups do not to appear to be separated by clear stratification variables such as landholdings or education, but rather more subtle differences such how they use this land and how well households are diversified, subjective grouping based on assumed household technological constraints would also misspecify household technologies in this sample. This is all supportive of the latent grouping specification in allocating heterogeneous households without a-priori stratification assumptions.

7 Livelihood Trajectories Results

Given that the livelihood technologies have been determined and significant heterogeneity has been observed, the fitted values can now be used to determine whether this heterogeneity matters in livelihood trajectories. The optimal trajectory models are chosen first, and then these trajectories are interpreted on their shape and levels. Finally, a number of specification tests are conducted to understand why conditional convergence is observed instead of structural poverty traps as described by Barrett et al. (2006).

7.1 Model Selection

The AIC and BIC statistics for the polynomial estimations and the control lag lengths are reported in tables (4) and (5). For most groups, the choice of polynomial beyond two or three orders has very little effect on the overall shape of the function for most observations, and rather accounts for a small number of higher fitted livelihoods. As the BIC tends to be more parsimonious compared to the AIC, the AIC is used as the first measure of the optimal polynomial order. The BIC is then used to allocate if the AIC cannot significantly differentiate between models. The lag lengths are then chosen for the optimal polynomial functions, where four groups have a lag length of four, and shorter lengths of two and three for the others. The results from the Durbin-Watson tests all conclude that there is no serial-correlation in the errors once these lag lengths are included. All the chosen model specifications can be found in table (6). The resulting trajectory equations are reported in the appendix in table (12), and are plotted in figure (4) for the latent grouping technology and figure (5) for the homogeneous technology. These plots are magnified to visually highlight the equilibrium points. Plots that account for the entire domain and range of the group livelihood levels can be found in the appendix as figures (7) and (8).

Table 4: AIC and BIC for Polynomial Trajectory Specifications

Cohort		Polynomial Order							
		1	2	3	4	5	6	7	8
Entire Sample	AIC	60,035	59,931	59,929	59,925	59,918	59,912	59,913	59,915
	BIC	60,053	59,954	59,958	59,961	59,960	59,959	59,966	59,974
Group 1	AIC	15,600	12,103	12,101	12,097	12,093	12,093	12,094	12,091
	BIC	15,614	12,120	12,122	12,122	12,123	12,127	12,132	12,134
Group 2	AIC	5,635	4,201	4,198	4,197	4,199	4,201	4,198	4,199
	BIC	5,645	4,214	4,214	4,217	4,222	4,227	4,227	4,231
Group 3	AIC	6,464	4,904	4,904	4,905	4,905	4,901	4,900	4,901
	BIC	6,475	4,918	4,921	4,926	4,929	4,928	4,931	4,936
Group 4	AIC	8,917	6,903	6,889	6,891	6,893	6,893	6,895	6,896
	BIC	8,929	6,918	6,908	6,913	6,919	6,923	6,929	6,934
Group 5	AIC	8,696	6,695	6,686	6,683	6,684	6,686	6,687	6,659
	BIC	8,708	6,710	6,705	6,705	6,710	6,715	6,720	6,695
Group 6	AIC	14,320	11,130	11,099	11,067	11,067	11,068	11,070	11,068
	BIC	14,333	11,147	11,120	11,093	11,096	11,102	11,108	11,110

Table 5: AIC and BIC for Control Lagged Livelihood Lengths in Trajectory Specifications

Cohort		Control Lag Length			
		No controls	$Livelihood_{t-2}$	$Livelihood_{t-3}$	$Livelihood_{t-4}$
Entire Sample	AIC	$59,929$	$55,111$	$50,319$	$45,543$
	BIC	$59,958$	$55,146$	$50,359$	$45,587$
Group 1	AIC	$12,093$	$12,088$	$12,074$	$12,073$
	BIC	$12,127$	$12,126$	$12,117$	$12,119$
Group 2	AIC	$4,198$	$4,193$	$4,195$	$4,195$
	BIC	$4,217$	$4,216$	$4,221$	$4,224$
Group 3	AIC	$4,904$	$4,707$	$4,539$	$4,415$
	BIC	$4,921$	$4,727$	$4,562$	$4,442$
Group 4	AIC	$6,890$	$6,845$	$6,805$	$6,764$
	BIC	$6,909$	$6,868$	$6,832$	$6,794$
Group 5	AIC	$6,684$	$6,618$	$6,575$	$6,530$
	BIC	$6,706$	$6,643$	$6,604$	$6,563$
Group 6	AIC	$11,067$	$11,039$	$11,031$	$11,022$
	BIC	$11,097$	$11,073$	$11,070$	$11,065$

Table 6: Results of Optimal Trajectory Functions

	Entire Sample	G1	G2	G3	G4	G5	G6
Observations	2,025	520	190	205	310	280	520
Polynomial Order	3	6	4	3	3	4	5
Control Lag Length	4	3	2	4	4	4	4
Durbin-Watson Test Statistic	1.982	1.992	2.315	1.952	1.998	1.886	2.032
Durbin-Watson p-value	0.338	0.435	0.981	0.355	0.462	0.153	0.605
Intersection	$9,892$	$26,091$	$21,518$	$7,684$	$20,054$	$32,131$	$6,659$

Note: Durbin Watson test conducted under an alternative hypothesis of autocorrelated residuals. The intersection point is the intersection between the 45° line and the trajectory function.

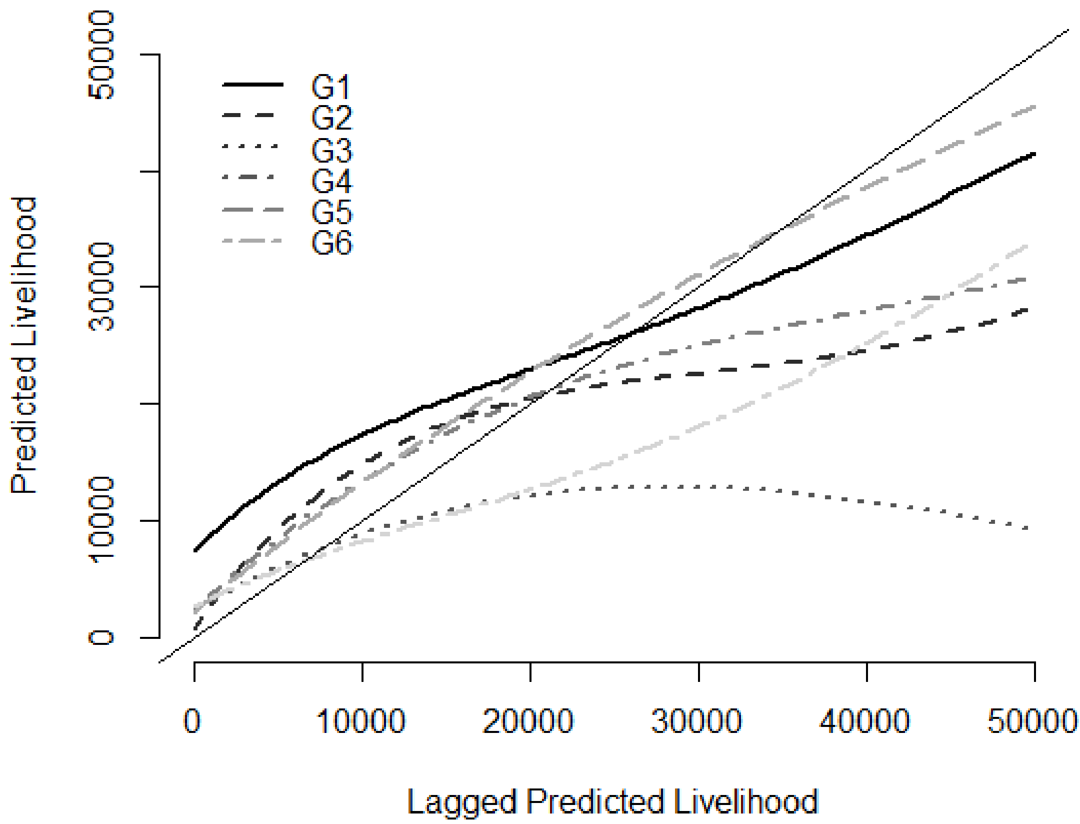

Figure 4: Trajectory Functions by Group

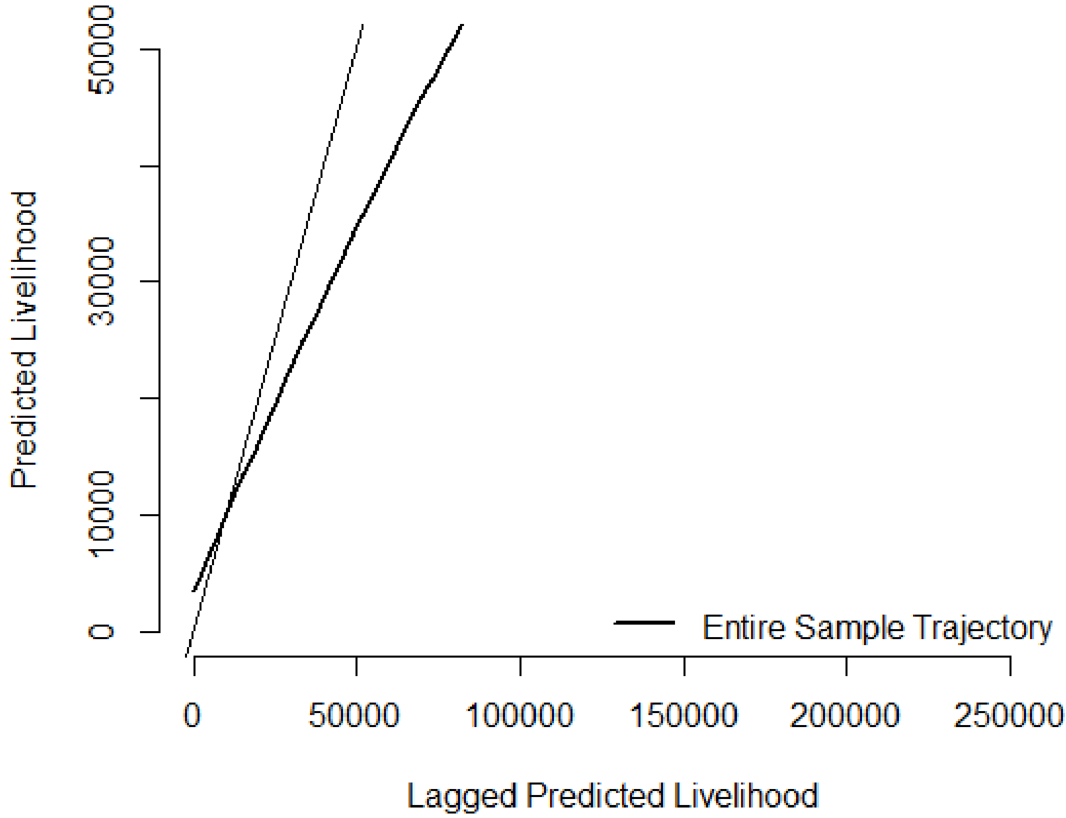

Figure 5: Homogeneous Technology Trajectory Function

7.2 Characteristics of Trajectory Functions

All trajectory functions converge to a conditional equilibrium points with the 45° line, but have large differences in the shape of trajectories. Groups Three and Six have a much faster decline in livelihoods towards their convergence levels, whereas groups One and Five have a much slower decline to a higher convergence level. These different rates in decline in livelihoods when below the 45° line shows differences in the groups in maintaining their asset stocks. For example, Group Six have a small but consistent income from farm labour and non-farm sources. They might be less likely to maintain a positive livelihood shock because they lack the wealth and income to be able to smooth asset stocks through negative shocks. Whereas Group Five, with larger asset holdings, may have enough wealth to delay the sale of assets during a negative shock. These differences in groups cannot be captured in a homogeneous technology specification as displayed in figure (5). If a policy such as a cash transfer is to be designed based on the results of a homogeneous trajectory analysis, the way in which households are able to maintain this shock is not understood. On the other hand, estimating livelihood trajectories with heterogeneous considerations, such as the latent grouping presented in this paper, can help differentiate how to target certain groups who would benefit the most from such a policy. In this case, heterogeneity does matter in estimating livelihood trajectories.

The intersection points between the 45° line and the trajectory functions are reported in table (6). As expected, the group with the highest median income and well diversified income sources, Group One, has the highest convergence level out of all the groups at just above 26,000 Rupees once adjusted by the requirements index per household. So households with an adjusted livelihood above this level may not be able to maintain these livelihood levels year by year, and those below are expected to increase their livelihoods. There is large differences in the convergence levels of these sub-technologies, with Group Six having the lowest livelihood convergence level at $6,659$ Rupees per the requirements index. This indicates that there is inequality in the technology sets available for households in the sample, where households are faced with constraints such as labour market opportunities in capturing a higher convergence equilibrium. Again, this is not observed in a homogeneous trajectory estimation, and yet is an important consideration when targeting development policy to alleviate constraints to higher convergence technologies.

What might be surprising is that Group Three, with significant asset holdings, has a much lower convergence level of 7,600 rupees. This would derive from the negative coefficients on land and livestock assets in the technology specification once household levels are included. A possible reason for this result could be from the falls in income from the years with productivity shocks, but a selection of households had increased their land and livestock inventories ex-ante to the lower productivity realisations. So even households at low levels of asset holdings, who accumulate assets over time in this group, are predicted to have a fall in their livelihood. Returning to the definition of livelihoods, which is the ability of a household to recover from shocks (Scoones 1998), a low convergence level is indicative of a group that has low capabilities of recovering from macroeconomic shocks. Although Group Three is vulnerable to macroeconomic shocks, the volatility in the time dummies also shows that they are able to recover reasonably well. So the low livelihood convergence level and observed income dynamics gives conflicting evidence as to how households in this group recover from shocks, and that an auto-regressive livelihood trajectory may poorly specify groups with high volatility. In scenarios such as this, true livelihood trajectories may be better identified in cases that account for the ability of a household to recover from a shock. Examples of such analysis include Carter et al. (2007) and Quisumbing & Baulch (2013), where positive and negative shocks to welfare such as drought, rainfall, dowry receipts and payments, illness and death are all included as controls in the construction of livelihood levels.

The homogeneous technology estimates a convergence much lower than most groups at a level of $9,892$ rupees. Only Group Three has a lower estimated convergence level amongst the latent groupings. What this might indicate is that a homogeneous technology is unable to identify what assets have the largest marginal contributions for each group. For example, the latent grouping results indicate an accumulation of irrigated land for a household in Group Six to have a large and significant effect on their livelihoods, as it may open up mobility and diversification opportunities for otherwise landless households. However, in a homogeneous estimation, households in Group One, who do not derive much from irrigated land as they source income from other areas, pulls the coefficient on irrigated land down towards zero. So any accumulation of irrigated land from a household in Group Six is under-reported by a homogeneous estimation. This effect is less pronounced in reverse, where other assets are less likely to be overestimated for some groups in the homogeneous technology. This is because there is a approximate truncation of asset elasticities in the latent grouping technologies around 0 or -0.05, meaning that assets are unlikely to ever have a major negative impact on livelihoods. The distance between each group's smaller coefficients and the respective homogeneous coefficients is shorter in this case, so the homogeneous estimation is less likely to over-emphasis the importance of assets for some groups. As a result, the homogeneous specification under-estimates livelihoods in a heterogeneous environment, resulting in a much lower convergence level than the average convergence of the latent group trajectories. Again, this shows that correctly identifying heterogeneity amongst households is an important consideration when making inference of livelihood dynamics, and that care should be taken interpreting previous literature that utilises a homogeneous technology for livelihoods.

7.3 Robustness Checks on Conditional Convergence

To ensure that the conclusions of conditional convergence are robust, alternative considerations as to possible drivers of these results need to be assessed. Firstly, the VDSA data is designed for a poor rural population, so this conditional convergence may not be representative of a conditional convergence when taken outside the context of the population represented in the sample. This single convergence equilibrium may only represent the lower equilibrium on trajectory, and that richer households not observed in the sample, but utilise the same sub-technology, may converge to an equilibrium much higher than what could be represented in the estimation. This can only be estimated or tested with panel data that has a greater proportion of wealthier households. As it has already been established that there is little concern of attrition in the sample at the upper end of wealth distributions, it can still be claimed that this convergence is still internally valid amongst poorer rural villages in India.

One other possible reason as to why we do not observe bifurcating asset dynamics could derive from asset holdings moving too slowly to show long run trajectories (Naschold 2012). This can be supported by papers such as Adato et al. (2006), where longer asset lags found the existence of multiple equilibria in the dynamic paths. Using the VDSA data however, Naschold (2012) failed to find multiple equilibria when using a 3-year lag, indicating that the one year lag is suitable to identify any asset-induced poverty traps. In the trajectory function presented in this paper, the inclusion of further lags are significant for all groups, and reduce the coefficients on the one-year lagged asset index. This most likely implies that further lags are positively correlated with the one-year lag and the resulting livelihood, inducing a positive bias in the coefficients if omitted from the model. Despite this relationship, further length lags have a diminishing upwards bias effect on the polynomial coefficients, and bifurcating asset dynamic are not observed for any model that includes the further lags or not. So even if assets have a delay in generating a livelihood, it does not change the shape of the trajectory towards one of bifurcating livelihoods as suggested by Barrett et al. (2006).

Another two issues raised by Naschold (2012) are that the assets obtained by a household may be shared through social networks, and households form endogenously through their asset holdings. The sharing of assets amongst social networks may be an issue, and is difficult to test without adequate social capital variables in the VDSA data. The VDSA data does include income from transfers, which represents gifts and transfers from other households. The count and median amount of these transfers are low relative to the sample size and total incomes, which partially alleviates the concern that poorer households may be receiving large transfers of kind that pulls results towards the mean. However, if possible, future studies should also include social capital variables in the livelihood regression, similar to that used by Moser & Felton (2009), as social capital has been argued to be an important asset for households to maintain a livelihood (Scoones 1998).

If the latter concern is true, this would mean that households that accumulate assets over time may grow their family, which dilutes the per-capita asset holdings each member of the household has. This is particularly true in India, where the existence of a marriage dowry (Field & Ambrus 2008) and preference for male children (Vogl 2013) can heavily influence the family composition of credit constrained households. To ensure this has not impacted the results, the correlation between the change in household size and the change in income is calculated to be 0.028. Without context or relative correlations with other studies, this correlation can give no conclusions about endogenous household compositions. However, what this might indicate is that changes in family composition occur too infrequently for this to have a strong effect on the livelihood regression and trajectories, and aligns with the findings of Naschold (2012) where trajectories are largely unaffected when using the household's total income and not income divided by household requirements.

8 Conclusion

This paper explores the commonly used livelihood regression as a tool for analysing poverty dynamics when alternative welfare measures are unreliable. Most studies assume a homogeneous process in which households derive a livelihood, which does not take into account how some households faces different opportunities and constraints in how they utilise assets to derive a livelihood. By redefining the livelihood regression as a livelihood generating technology, this paper shows how insights from production literature allow for greater considerations in how households may differ in their technology sets.

A latent grouping approach allows for identification of these technology sets without making assumptions of how these technology sets are defined. Varying magnitudes of both the elasticity coefficients and the efficiency constants show that assumptions of homogeneity in constructing a livelihood index are unreasonable and that livelihoods estimated in the previous literature may be misspecified. An analysis into the characteristics of these groups show that sub-technologies are associated with a diversification of income sources, vulnerability to macroeconomic shocks and access to collective technologies. These characteristics are subtle, and support the case that subjective stratification is a unlikely to allocate households into the correct technology sets.

Finally, trajectory functions are estimated to test whether using heterogeneous livelihood technologies matter in the analysis of livelihood trajectories. All groups are found to converge to conditional livelihood levels, albeit with varying trajectory shapes and levels. These trajectories show that some groups converge to levels more quickly than others, indicating that the are less likely to maintain a positive livelihood shock. A homogeneous trajectory is unable to identify these differences, and is important from a policy perspective as it cannot determine the groups that are most likely to maintain asset transfers over time. Finally, the difference in levels show that some sub-technologies have much smaller livelihood outcomes compared to others, and supports previous studies in rural India that diversification of income sources allows households to achieve higher livelihood levels. This shows an inequality between the outcomes from different technology sets, one of which a homogeneous trajectory estimation cannot identify. Overall, this paper has highlighted that the previous literature has under-appreciated heterogeneity in livelihood trajectories, and care should be taken interpreting results that use a homogeneous livelihood estimations.

References

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

Alderman, H., Behrman, J. R., Kohler, H.-P., Maluccio, J. A. & Watkins, S. C. (2001), 'Attrition in longitudinal household survey data: some tests for three developing-country samples', *Demographic research* **5**, 79–124.

Amsler, C., O'Donnell, C. J. & Schmidt, P. (2017), 'Stochastic metafrontiers', *Econometric Reviews* **36**(6-9), 1007–1020.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B. & Udry, C. (2015), 'A multifaceted program causes lasting progress for the very poor: Evidence from six countries', *Science* **348**(6236), 1260799.

Banerjee, A. V. & Newman, A. F. (1993), 'Occupational choice and the process of development', *Journal of political economy* **101**(2), 274–298.

Barrett, C. B. & Carter, M. R. (2013), 'The economics of poverty traps and persistent poverty: empirical and policy implications', *The Journal of Development Studies* **49**(7), 976–990.

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J. C., Rasambainarivo, J. & Wangila, J. (2006), 'Welfare dynamics in rural kenya and madagascar', *The Journal of Development Studies* **42**(2), 248–277.

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

Carter, M. R., Little, P. D., Mogues, T. & Negatu, W. (2007), 'Poverty traps and natural disasters in ethiopia and honduras', *World development* **35**(5), 835–856.

Carter, M. R. & Lybbert, T. J. (2012), 'Consumption versus asset smoothing: testing the implications of poverty trap theory in Burkina Faso', *Journal of Development Economics* **99**(2), 255–264.

Carter, M. R. & May, J. (2001), 'One kind of freedom: Poverty dynamics in post-apartheid South Africa', *World development* **29**(12), 1987–2006.

Chaudhuri, S. & Ravallion, M. (1994), 'How well do static indicators identify the chronically poor?', *Journal of Public Economics* **53**(3), 367–394.

Dasgupta, P. & Ray, D. (1986), 'Inequality as a determinant of malnutrition and unemployment: Theory', *The Economic Journal* **96**(384), 1011–1034.

Deaton, A. & Zaidi, S. (2002), *Guidelines for constructing consumption aggregates for welfare analysis*, Vol. 135, World Bank Publications.

Deb, U., Bantilan, C. & Anupama, G. (2014), Dynamics of rural livelihoods and poverty in SAT India, Research Bulletin No. 26, International Crops Research Institute for the Semi-Arid Tropics, Telegana, India.

Färe, R., Grosskopf, S., Lovell, C. K. et al. (1994), *Production Frontiers*, Cambridge University Press.

Field, E. & Ambrus, A. (2008), 'Early marriage, age of menarche, and female schooling attainment in bangladesh', *Journal of political Economy* **116**(5), 881–930.

Filmer, D. & Scott, K. (2012), 'Assessing asset indices', *Demography* **49**(1), 359–392.

Galor, O. & Zeira, J. (1993), 'Income distribution and macroeconomics', *The review of economic studies* **60**(1), 35–52.

Gang, I. N., Sen, K. & Yun, M.-S. (2008), ‘Poverty in rural india: caste and tribe’, *Review of Income and Wealth* **54**(1), 50–70.

Giesbert, L. & Schindler, K. (2012), ‘Assets, shocks, and poverty traps in rural Mozambique’, *World Development* **40**(8), 1594–1609.

Greene, W. (2005), ‘Reconsidering heterogeneity in panel data estimators of the stochastic frontier model’, *Journal of Econometrics* **126**(2), 269–303.

Jalan, J. & Ravallion, M. (2004), ‘Household income dynamics in rural china’, *Insurance against poverty* pp. 108–124.

Kwak, S. & Smith, S. C. (2013), ‘Regional agricultural endowments and shifts of poverty trap equilibria: Evidence from Ethiopian panel data’, *The Journal of Development Studies* **49**(7), 955–975.

Latruffe, L., Bravo-Ureta, B. E., Carpentier, A., Desjeux, Y. & Moreira, V. H. (2017), ‘Subsidies and technical efficiency in agriculture: Evidence from European dairy farms’, *American Journal of Agricultural Economics* **99**(3), 783–799.

Lin, B. & Du, K. (2013), ‘Technology gap and China’s regional energy efficiency: a parametric metafrontier approach’, *Energy Economics* **40**, 529–536.

Michelson, H., Muñiz, M. & DeRosa, K. (2013), ‘Measuring socio-economic status in the millennium villages: the role of asset index choice’, *The Journal of Development Studies* **49**(7), 917–935.

Moreira, V. H. & Bravo-Ureta, B. E. (2010), ‘Technical efficiency and metatechnology ratios for dairy farms in three southern cone countries: a stochastic meta-frontier model’, *Journal of Productivity Analysis* **33**(1), 33–45.

Moser, C. & Felton, A. (2009), 'The construction of an asset index', *Poverty dynamics: interdisciplinary perspectives* pp. 102–127.

Moser, C. O. (1998), 'The asset vulnerability framework: reassessing urban poverty reduction strategies', *World development* **26**(1), 1–19.

Mosse, D. (2018), 'Caste and development: Contemporary perspectives on a structure of discrimination and advantage', *World Development* **110**, 422–436.

Naschold, F. (2012), '"the poor stay poor": Household asset poverty traps in rural semi-arid India', *World Development* **40**(10), 2033–2043.

Naschold, F. (2013), 'Welfare dynamics in Pakistan and Ethiopia–does the estimation method matter?', *The Journal of Development Studies* **49**(7), 936–954.

O'Donnell, C. (2016), 'Using information about technologies, markets and firm behaviour to decompose a proper productivity index', *Journal of Econometrics* **190**(2), 328–340.

O'Donnell, C. J., Rao, D. P. & Battese, G. E. (2008), 'Metafrontier frameworks for the study of firm-level efficiencies and technology ratios', *Empirical economics* **34**(2), 231–255.

Orea, L. & Kumbhakar, S. C. (2004), 'Efficiency measurement using a latent class stochastic frontier model', *Empirical economics* **29**(1), 169–183.

Quisumbing, A. R. & Baulch, B. (2013), 'Assets and poverty traps in rural bangladesh', *The Journal of Development Studies* **49**(7), 898–916.

Rao, K. (2008), 'Changes in dry land agriculture in the semi-arid tropics of India, 1975–2004', *The European Journal of Development Research* **20**(4), 562–578.

Romer, P. M. (1990), 'Endogenous technological change', *Journal of political Economy* **98**(5, Part 2), S71–S102.

Ryan, J. G., Bidinger, P. D., Rao, N. P. & Pushpamma, P. (1984), 'The determinants of individual diets and nutritional status in six villages of southern india', *ICRISAT Research Bulletin no 7* .

Scoones, I. (1998), 'Sustainable rural livelihoods: a framework for analysis'.

Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. (2016), 'mclust 5: Clustering, classification and density estimation using gaussian finite mixture models', *The R journal* **8**(1), 289.

Solow, R. M. (1957), 'Technical change and the aggregate production function', *The review of Economics and Statistics* **39**(3), 312–320.

Thorat, A., Vanneman, R., Desai, S. & Dubey, A. (2017), 'Escaping and falling into poverty in india today', *World development* **93**, 413–426.

Thorat, S. & Neuman, K. S. (2012), *Blocked by caste: economic discrimination in modern India*, Oxford University Press.

Vogl, T. S. (2013), 'Marriage institutions and sibling competition: Evidence from south asia', *The Quarterly journal of economics* **128**(3), 1017–1072.

Walker, T. S. & Ryan, J. G. (1990), *Village and household economics in India's semi-arid tropics*, Johns Hopkins University Press.

Wang, Q., Zhao, Z., Zhou, P. & Zhou, D. (2013), 'Energy efficiency and production technology heterogeneity in China: a meta-frontier dea approach', *Economic Modelling* **35**, 283–289.

World Bank Group (2016), *Poverty and shared prosperity 2016: taking on inequality*, World Bank Publications.

Appendix

Table 7: Cobb Douglas Homogeneous Livelihood Regression

Dep. Var:	*OLS*	*Fixed Effects Estimates*			*Random Effects Estimates*		
INCOME		Individual	Time	Two-Way	Individual	Time	Two-Way
Irrigated Land	−0.165***	0.012	−0.141***	0.134**	−0.108**	−0.144***	−0.025
	(0.030)	(0.056)	(0.030)	(0.054)	(0.042)	(0.030)	(0.054)
Dryland	−0.116***	0.113*	−0.116***	0.118**	−0.033	−0.117***	−0.010
	(0.038)	(0.061)	(0.037)	(0.057)	(0.049)	(0.037)	(0.061)
Large Livestock	−0.018***	0.003	−0.019***	0.005	−0.007	−0.018***	−0.005
	(0.004)	(0.006)	(0.004)	(0.006)	(0.005)	(0.004)	(0.007)
Small Livestock	0.004	−0.007	0.006	−0.004	−0.003	0.006	−0.001
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.007)
Durables	0.248***	0.203***	0.233***	0.103***	0.216***	0.235***	0.154***
	(0.016)	(0.017)	(0.018)	(0.022)	(0.016)	(0.018)	(0.025)
Buildings	0.055***	0.056***	0.046***	0.025	0.059***	0.047***	0.036*
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.020)
Farm Equipment	0.051***	0.058***	0.054***	0.053***	0.049***	0.054***	0.051***
	(0.012)	(0.015)	(0.011)	(0.014)	(0.014)	(0.011)	(0.017)
Education	0.090***	0.201***	0.090***	0.167***	0.117***	0.090***	0.110***
	(0.023)	(0.047)	(0.022)	(0.045)	(0.034)	(0.022)	(0.042)
Constant	7.247***	0.038**	7.459***	−0.104*	7.347***	7.415***	8.122***
	(0.153)	(0.015)	(0.182)	(0.061)	(0.162)	(0.198)	(0.297)
Observations	2,900	2,900	2,900	2,900	2,900	2,900	2,900
R^2	0.194	0.150	0.259	0.257	0.160	0.119	0.173
Adjusted R^2	0.191	0.147	0.254	0.251	0.158	0.116	0.171
Hausman Test χ^2					40.125	1.4847	43.944
LM Test χ^2					836.53	1713.1	2549.6
AIC				6665.1			
BIC				6802.5			

Note: Time coefficients are omitted for brevity. Hausman test statistics under random models are associated with corresponding fixed effect model. Hausman test results reject random individual and twoway effects, and cannot reject random time effects. Breusch-Pagan LM tests reject all random effects, meaning two-way fixed effects is preferred. AIC and BIC statistics only included for two-way fixed effects as other models rejected. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 8: Translog Homogeneous Livelihood Regression

Dep. Var:	*OLS*	*Fixed Effects Estimates*			*Random Effects Estimates*		
INCOME		Indiv	Time	Two-Way	Indiv	Time	Two-Way
IRRIGATED	0.082	0.065	0.097	0.152***	0.043	0.082	−0.744**
	(0.275)	(0.060)	(0.264)	(0.058)	(0.290)	(0.275)	(0.308)
DRYLAND	−0.409	0.117*	−0.504	0.122**	−0.262	−0.409	0.043
	(0.368)	(0.064)	(0.354)	(0.061)	(0.366)	(0.368)	(0.461)
LRG_LIVESTOCK	−0.127***	0.005	−0.138***	0.006	−0.105**	−0.127***	−0.149***
	(0.043)	(0.006)	(0.042)	(0.006)	(0.045)	(0.043)	(0.055)
SML_LIVESTOCK	0.064	−0.003	0.081*	−0.002	0.050	0.064	0.140**
	(0.049)	(0.006)	(0.047)	(0.006)	(0.048)	(0.049)	(0.059)
BUILDINGS	−0.137*	0.094***	−0.143*	0.044*	−0.196**	−0.137*	−0.107
	(0.082)	(0.023)	(0.079)	(0.025)	(0.079)	(0.082)	(0.100)
DURABLES	−0.215**	0.215***	−0.191**	0.146***	−0.139	−0.215**	−0.065
	(0.092)	(0.019)	(0.090)	(0.024)	(0.088)	(0.092)	(0.109)
FARM_EQUIP	−0.155	0.061***	−0.124	0.059***	−0.090	−0.155	−0.302**
	(0.098)	(0.016)	(0.095)	(0.015)	(0.095)	(0.098)	(0.119)
EDUC	−0.216	0.168***	−0.229	0.145***	−0.370*	−0.216	0.086
	(0.215)	(0.050)	(0.207)	(0.047)	(0.224)	(0.215)	(0.276)
IRRIGATEDxDRYLAND	0.228***	−0.056	0.235***	−0.108	0.144**	0.228***	0.0001
	(0.053)	(0.146)	(0.052)	(0.139)	(0.061)	(0.053)	(0.054)
IRRIGATEDxLRG_LIVESTOCK	−0.008	−0.028	−0.007	−0.026	−0.009	−0.008	−0.020**
	(0.008)	(0.021)	(0.007)	(0.019)	(0.008)	(0.008)	(0.008)
IRRIGATEDxSML_LIVESTOCK	0.008	−0.006	0.008	−0.021	−0.003	0.008	0.011
	(0.009)	(0.022)	(0.009)	(0.021)	(0.009)	(0.009)	(0.009)
IRRIGATEDxBUILDINGS	−0.025	0.025	−0.026	0.036	−0.024	−0.025	0.018
	(0.026)	(0.063)	(0.025)	(0.060)	(0.026)	(0.026)	(0.034)
IRRIGATEDxDURABLES	0.066**	0.044	0.069**	0.025	0.050*	0.066**	0.059*
	(0.029)	(0.065)	(0.028)	(0.061)	(0.029)	(0.029)	(0.033)
IRRIGATEDxFARM_EQUIP	−0.043*	−0.028	−0.042*	0.032	−0.023	−0.043*	−0.028
	(0.023)	(0.063)	(0.022)	(0.060)	(0.025)	(0.023)	(0.024)
IRRIGATEDxEDUC	−0.123***	0.075	−0.135***	0.077	−0.098**	−0.123***	−0.009
	(0.037)	(0.141)	(0.036)	(0.133)	(0.043)	(0.037)	(0.041)
DRYLANDxLRG_LIVESTOCK	−0.010	−0.038*	−0.011	−0.030	0.009	−0.010	0.002
	(0.010)	(0.023)	(0.009)	(0.022)	(0.011)	(0.010)	(0.012)

Table 9: Translog Homogeneous Livelihood Regression Continued

Dep. Var:	*OLS*	*Fixed Effects Estimates*			*Random Effects Estimates*		
INCOME		Indiv	Time	Two-Way	Indiv	Time	Two-Way
DRYLANDxSML_LIVESTOCK	0.036***	0.012	0.032***	−0.001	0.024*	0.036***	0.041***
	(0.012)	(0.023)	(0.012)	(0.022)	(0.012)	(0.012)	(0.015)
DRYLANDxBUILDINGS	−0.013	0.031	−0.003	0.020	−0.015	−0.013	−0.069
	(0.041)	(0.074)	(0.039)	(0.070)	(0.039)	(0.041)	(0.052)
DRYLANDxDURABLES	0.003	−0.010	0.004	−0.001	−0.016	0.003	0.019
	(0.038)	(0.071)	(0.037)	(0.067)	(0.037)	(0.038)	(0.047)
DRYLANDxFARM_EQUIP	−0.029	0.019	−0.024	0.032	−0.003	−0.029	0.004
	(0.025)	(0.059)	(0.024)	(0.056)	(0.026)	(0.025)	(0.032)
DRYLANDxEDUC	−0.068	0.078	−0.094*	0.089	0.007	−0.068	−0.032
	(0.053)	(0.173)	(0.050)	(0.164)	(0.056)	(0.053)	(0.067)
LRG_LIVESTOCKxSML_LIVESTOCK	0.003**	0.001	0.003**	0.001	0.003**	0.003**	0.005***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
LRG_LIVESTOCKxBUILDINGS	−0.002	−0.003	−0.001	−0.006	−0.005	−0.002	−0.0001
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.004)	(0.005)
LRG_LIVESTOCKxDURABLES	0.002	0.004	0.004	0.010	0.003	0.002	0.005
	(0.005)	(0.007)	(0.004)	(0.007)	(0.004)	(0.005)	(0.006)
LRG_LIVESTOCKxFARM_EQUIP	0.003	−0.001	0.002	−0.004	0.001	0.003	0.004
	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.003)	(0.004)
LRG_LIVESTOCKxEDUC	0.010	−0.026	0.011*	−0.024	0.008	0.010	0.006
	(0.006)	(0.018)	(0.006)	(0.017)	(0.007)	(0.006)	(0.008)
SML_LIVESTOCKxBUILDINGS	0.005	−0.003	0.004	−0.003	0.001	0.005	0.0001
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
SML_LIVESTOCKxDURABLES	−0.004	0.008	−0.004	0.008	−0.001	−0.004	−0.009
	(0.005)	(0.007)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)
SML_LIVESTOCKxFARM_EQUIP	−0.009**	−0.005	−0.008**	−0.005	−0.008**	−0.009**	−0.005
	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)
SML_LIVESTOCKxEDUC	−0.007	−0.011	−0.008	−0.007	−0.010	−0.007	−0.024***
	(0.006)	(0.018)	(0.006)	(0.017)	(0.007)	(0.006)	(0.009)
BUILDINGSxDURABLES	0.0001	−0.001	0.004	−0.003	0.004	0.0001	−0.007
	(0.010)	(0.011)	(0.009)	(0.010)	(0.009)	(0.010)	(0.011)
BUILDINGSxFARM_EQUIP	0.009	−0.005	0.007	0.001	0.012	0.009	0.020
	(0.012)	(0.014)	(0.011)	(0.014)	(0.011)	(0.012)	(0.014)

Table 10: Translog Homogeneous Livelihood Regression Continued

Dep. Var:	*OLS*	*Fixed Effects Estimates*			*Random Effects Estimates*		
INCOME		Indiv	Time	Two-Way	Indiv	Time	Two-Way
BUILDINGSxEDUC	0.003	0.041	0.009	0.035	0.015	0.003	−0.013
	(0.022)	(0.049)	(0.021)	(0.046)	(0.022)	(0.022)	(0.027)
DURABLESxFARM_EQUIP	−0.003	0.018	−0.006	0.011	−0.005	−0.003	−0.0005
	(0.011)	(0.015)	(0.011)	(0.014)	(0.011)	(0.011)	(0.014)
DURABLESxEDUC	0.050**	−0.058	0.044**	−0.050	0.034*	0.050**	−0.001
	(0.020)	(0.051)	(0.020)	(0.048)	(0.020)	(0.020)	(0.025)
FARM_EQUIPxEDUC	−0.033**	−0.047	−0.029*	−0.054	−0.019	−0.033**	0.042**
	(0.016)	(0.046)	(0.015)	(0.043)	(0.017)	(0.016)	(0.020)
IRRIGATED^2	−0.120***	−0.035	−0.124***	0.017	−0.054	−0.120***	0.007
	(0.041)	(0.084)	(0.040)	(0.079)	(0.049)	(0.041)	(0.053)
DRYLAND^2	0.283***	−0.161	0.283***	−0.164*	0.127**	0.283***	0.204**
	(0.057)	(0.099)	(0.055)	(0.094)	(0.062)	(0.057)	(0.079)
LRG_LIVESTOCK^2	0.008**	−0.0002	0.007**	0.0003	0.010***	0.008**	0.006
	(0.003)	(0.001)	(0.003)	(0.001)	(0.004)	(0.003)	(0.004)
SML_LIVESTOCK^2	−0.003	−0.003**	−0.004*	−0.003**	−0.0004	−0.003	−0.002
	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.003)	(0.004)
BUILDINGS^2	0.009**	0.008**	0.007**	0.003	0.010**	0.009**	0.010**
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
DURABLES^2	0.017**	0.023**	0.014**	0.015*	0.011	0.017**	0.016*
	(0.007)	(0.009)	(0.007)	(0.009)	(0.007)	(0.007)	(0.009)
FARM_EQUIP^2	0.019***	0.003	0.019***	0.003	0.009**	0.019***	0.004
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)
EDUC_PC^2	0.034	0.017	0.031	−0.011	0.059*	0.034	−0.017
	(0.025)	(0.060)	(0.024)	(0.057)	(0.032)	(0.025)	(0.038)
Constant	11.197***	0.028	11.120***	−0.060	10.983***	11.197***	10.435***
	(0.621)	(0.022)	(0.601)	(0.066)	(0.605)	(0.621)	(0.759)
Observations	2,900	2,900	2,900	2,900	2,900	2,900	2,900
R^2	0.256	0.174	0.319	0.270	0.203	0.256	0.187
Adjusted R^2	0.244	0.161	0.305	0.255	0.191	0.244	0.172
Hausman Test χ^2					254.74	113.41	121.34
LM Test χ^2					496.74	130.87	2381.6
AIC				6785.0			
BIC				7038.5			

Note: Time coefficients are omitted for brevity. Hausman test statistics under random models are associated with corresponding fixed effect model. Hausman test results reject random effects for all models. Breusch-Pagan LM tests reject all random effects, meaning two-way fixed effects is preferred. AIC and BIC statistics only included for two-way fixed effects as other models rejected.
*p<0.1; **p<0.05; ***p<0.01

Table 11: Time Dummy Estimates for Latent Grouping

	Entire Sample	*Latent Group Estimates* (1)	(2)	(3)	(4)	(5)	(6)
2002	0.001	0.054	−0.063	−0.059	−0.026	0.095	0.054
	(0.075)	(0.093)	(0.274)	(0.334)	(0.130)	(0.210)	(0.071)
2003	0.070	0.191**	−0.030	0.025	0.180	−0.043	0.016
	(0.076)	(0.094)	(0.274)	(0.361)	(0.131)	(0.214)	(0.072)
2004	0.063	0.197**	0.182	0.044	0.125	−0.062	−0.098
	(0.076)	(0.094)	(0.271)	(0.338)	(0.132)	(0.210)	(0.072)
2005	−0.511***	0.080	−0.013	−1.816***	−0.655***	−1.566***	−0.103
	(0.078)	(0.097)	(0.276)	(0.359)	(0.135)	(0.215)	(0.074)
6200	−0.438***	0.366***	0.069	−3.697***	−0.308**	−0.804***	0.069
	(0.079)	(0.099)	(0.287)	(0.358)	(0.134)	(0.216)	(0.078)
2007	0.640***	0.672***	0.533*	0.695*	0.889***	0.708***	0.396***
	(0.080)	(0.102)	(0.286)	(0.357)	(0.136)	(0.218)	(0.080)
2008	−0.060	0.250**	0.340	−1.778***	0.308**	−0.769***	0.337***
	(0.082)	(0.105)	(0.295)	(0.366)	(0.139)	(0.226)	(0.086)
2009	0.283***	0.590***	0.427	−0.843**	0.435***	0.020	0.393***
	(0.085)	(0.110)	(0.311)	(0.383)	(0.141)	(0.231)	(0.089)
2010	0.280***	0.330***	0.402	−0.083	0.448***	0.479**	0.073
	(0.087)	(0.113)	(0.313)	(0.402)	(0.146)	(0.236)	(0.092)
2011	0.375***	0.384***	0.226	−0.335	0.770***	0.665***	0.235**
	(0.089)	(0.117)	(0.328)	(0.393)	(0.150)	(0.244)	(0.096)
2012	0.584***	0.527***	0.441	−0.079	0.958***	0.900***	0.473***
	(0.094)	(0.125)	(0.340)	(0.426)	(0.157)	(0.258)	(0.102)
2013	0.460***	0.361***	−0.647*	0.326	1.021***	0.609**	0.449***
	(0.096)	(0.128)	(0.339)	(0.438)	(0.160)	(0.266)	(0.109)
2014	0.218**	0.301**	−1.095***	−0.206	0.620***	0.663**	0.171
	(0.098)	(0.132)	(0.347)	(0.442)	(0.162)	(0.268)	(0.115)
Observations	2,900	728	266	328	446	404	728
R^2	0.257	0.482	0.282	0.500	0.514	0.512	0.494
Adjusted R^2	0.251	0.467	0.220	0.466	0.490	0.486	0.479

Note: 2001 is the base year *p<0.1; **p<0.05; ***p<0.01

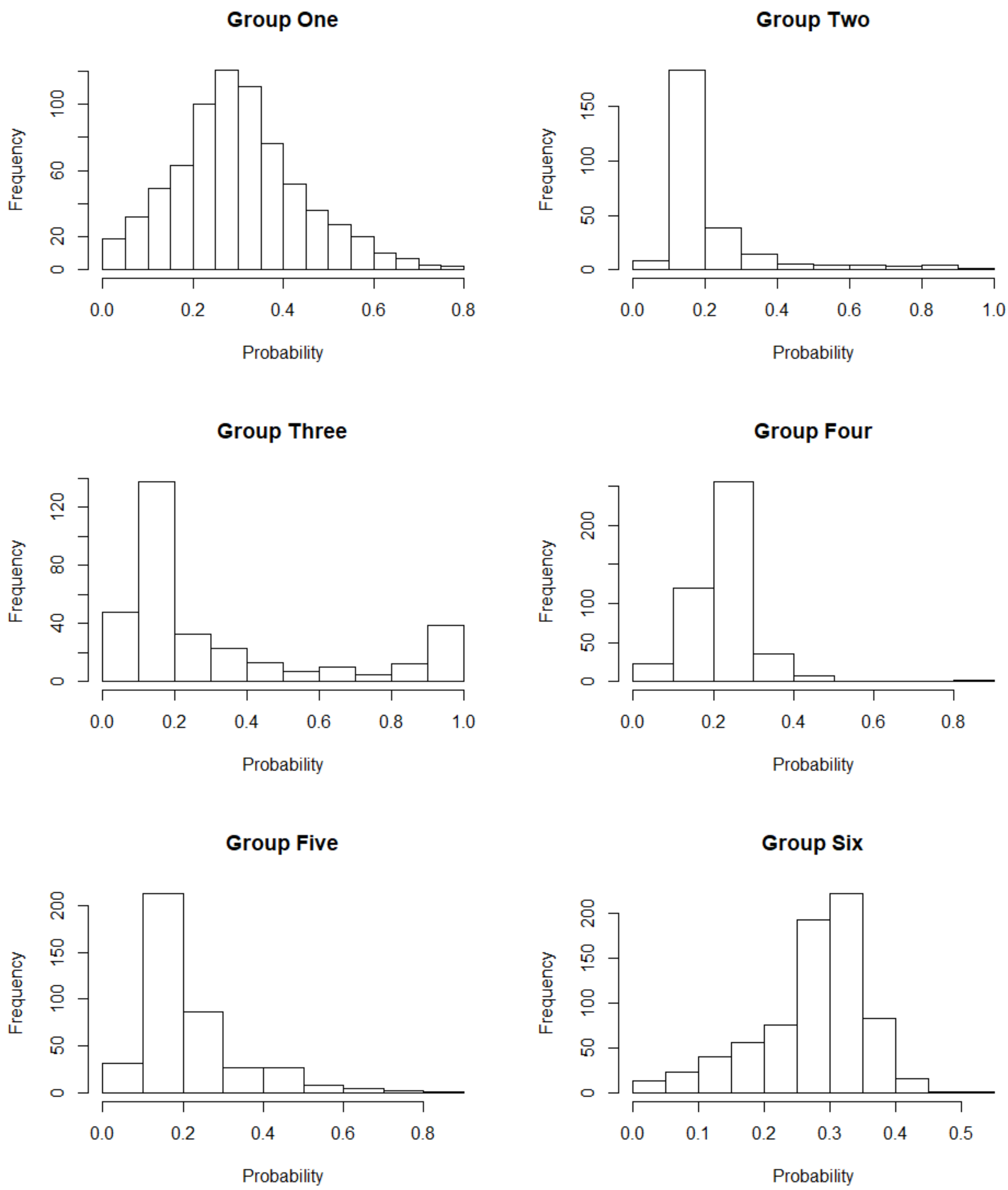

Figure 6: Histograms of Probability of Belonging in Group Allocated To in Time t

Table 12: Trajectory Function Results

	Dependent variable: Livelihood$_t$						
		Latent Groups					
	Entire Sample	(G1)	(G2)	(G3)	(G4)	(G5)	(G6)
Livelihood$_{t-1}$	0.65***	1.43**	2.01***	0.84***	1.35***	1.18***	0.702**
	(0.056)	(0.680)	(0.485)	(0.249)	(0.253)	(0.287)	(0.331)
Livelihood$^2_{t-1}$	−4.51e-07	− 5.83e-5	−7.24e-5**	−1.92e-5*	−2.63e-5***	−9.42e-6	−2.07e-5
	(9.61e-7)	(4.51e-5)	(3.06e-5)	(1.03e-5)	(8.52e-6)	(6.31e-6)	(3.30e-5)
Livelihood$^3_{t-1}$	− 4.02e-12	1.60e-9	1.10e-9*	1.13e-10	2.12e-10***	6.73e-11	5.73e-10
	(3.57e-12)	(1.10e-9)	(6.28e-10)	(9.02e-11)	(7.04e-11)	(4.25e-11)	(1.10e-9)
Livelihood$^4_{t-1}$		−2.15e-14*	−5.97e-15			−1.60e-16*	−3.59e-15
		(1.23e-14)	(3.87e-15)			(8.58e-17)	(1.44e-14)
Livelihood$^5_{t-1}$		1.26e-19**					−1.26e-19
		(1.2e-19)					(6.4e-20)
Livelihood$^6_{t-1}$		−2.71e-25**					
		(1.21e-25)					
Livelihood$_{t-2}$	0.135***	0.022	0.1078**	0.024	0.118	0.309***	0.15**
	(0.026)	(0.055)	(0.043)	(0.073)	(0.085)	(0.066)	(0.065)
Livelihood$_{t-3}$	0.168***	0.21***		0.32***	−0.097	0.056	0.037
	(0.029)	(0.05)		(0.06)	(0.093)	(0.07)	(0.083)
Livelihood$_{t-4}$	−0.003			0.087	0.109	−0.095	0.251***
	(0.028)			(0.073)	(0.088)	(0.074)	(0.075)
Constant	3,461.5***	7,366.8***	814.3	2,331.4**	2,328.7***	2,201.3	2,711.4***
	(545.6)	(1,930.6)	(1,497.4)	(1,083.5)	(891.2)	(2,200.4)	(568.7)
Observations	2,025	520	190	205	310	280	520
Polynomial Order	3	6	4	3	3	4	5
Control Lag Length	4	3	2	4	4	4	4
R^2	0.5097	0.5168	0.4743	0.2938	0.4741	0.6584	0.6468
Adjusted R^2	0.5083	0.5093	0.4600	0.2724	0.4637	0.6496	0.6412
Durbin-Watson Statistic	1.982	1.992	2.315	1.952	1.998	1.886	2.032
Durbin-Watson p-value	0.338	0.435	0.981	0.355	0.462	0.153	0.605

Coefficients rounded for brevity. Entire sample represents the homogeneous estimation, and Latent Groups for the groupings derived in the technology estimations. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

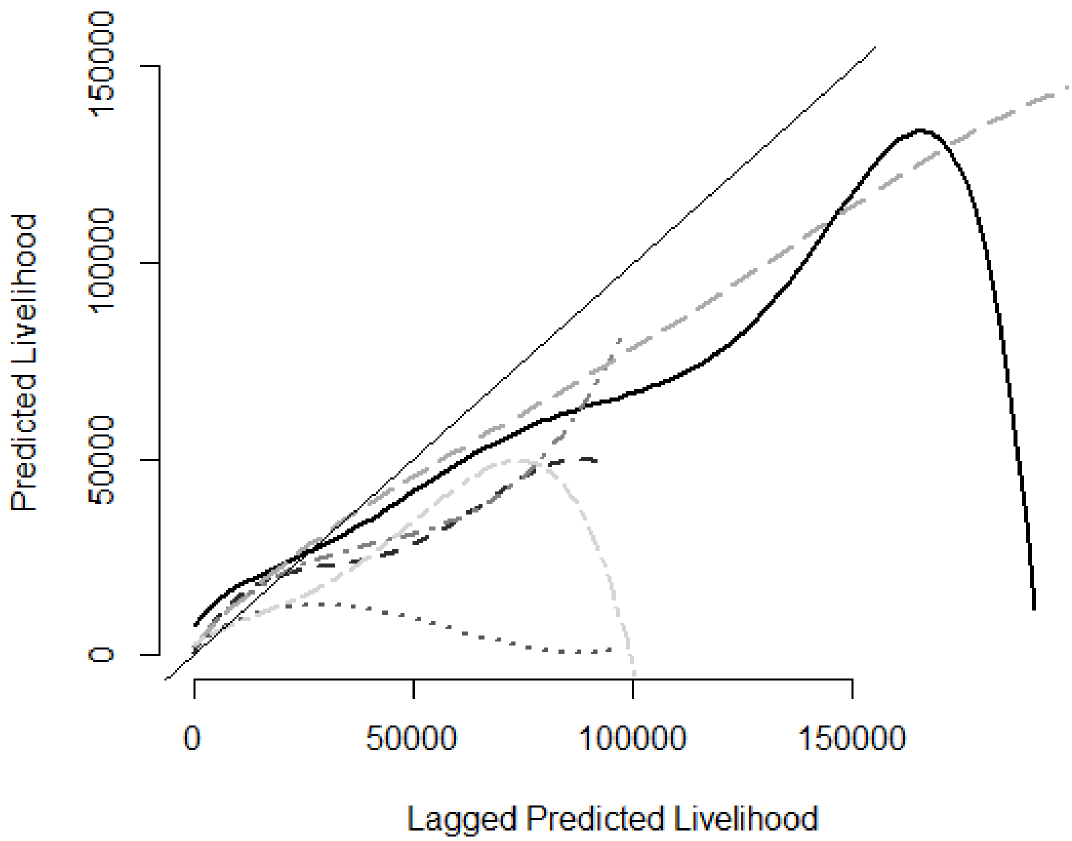

Figure 7: Trajectory Function by Group, estimated for entire domains and range of estimated livelihoods

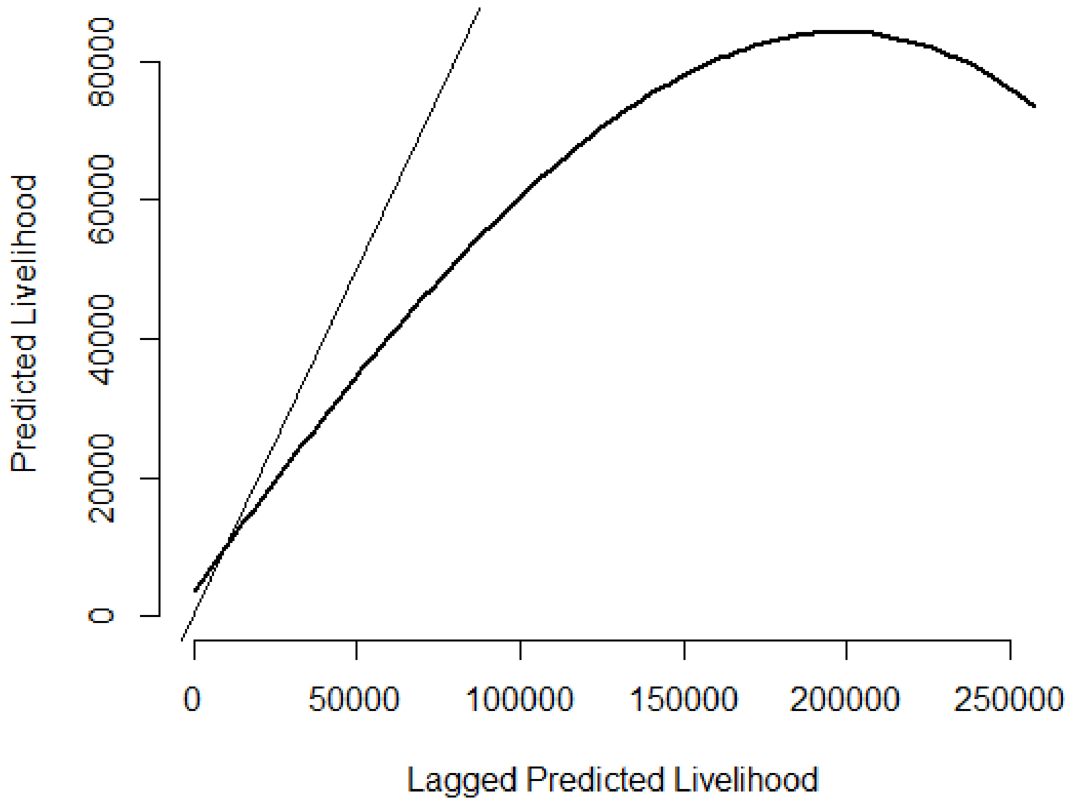

Figure 8: Homogeneous Technology Trajectory Function, estimated for entire domains and range of estimated livelihoods