CMI WORKING PAPER

WP 2012: 11

Poverty dynamics in rural Orissa

Transitions in assets and occupations over generations

Magnus Hatlebakk

CMI CHR. MICHELSEN INSTITUTE

Chr. Michelsen Institute (CMI) is an independent, non-profit research institution and a major international centre in policy-oriented and applied development research. Focus is on development and human rights issues and on international conditions that affect such issues. The geographical focus is Sub-Saharan Africa, Southern and Central Asia, the Middle East and Latin America.

CMI combines applied and theoretical research. CMI research intends to assist policy formulation, improve the basis for decision-making and promote public debate on international development issues.

Poverty dynamics in rural Orissa:

Transitions in assets and occupations over generations

Magnus Hatlebakk (CMI)*

WP 2012: 11

December 2012

CHR. MICHELSEN INSTITUTE

* Data was collected by RAC, Bhubaneswar, Orissa. The research is part of a larger collaboration with Basanta Pradhan, IEG-Delhi, that is sponsored by the Norwegian Embassy in Delhi. The paper was completed during a visit to the Dyson School of Applied Economics and Management at Cornell University.

Project number
29025

Project title
Inclusive growth in India

JEL-classification
D310, D910, O120

Contents

Abstract

We investigate whether historic land distribution determines stagnation or development of Indian villages. The empirical analysis is motivated by the Banerjee and Newman (1993) model of occupational choice and economic development. Family histories are collected for a random sample of 800 households. Households are classified into economic categories according to the assets-occupations mix at present and at grandfather's time. Transitions are described, and for a remote district explained, by the historic village land distribution. We also investigate the role of social identity, and find that scheduled tribes are more likely trapped in poverty than scheduled castes.

1. Introduction

There are mechanisms at both the household and economy level that contribute to the long-term persistence of poverty (Barrett and Carter, 2012). Households can be in a liquidity, or credit, trap where they do not get loans to finance investments or insurance because they do not have capital in the first place that can be used as collateral to raise additional capital. Lack of insurance can lead to conservative choices where the households stick to traditional technology in stead of experimenting with potentially more effective ones, and modern technology may require capital that they cannot afford.

Poverty can also explain preferences that in turn lead to conservative behavior that reproduces poverty. As the permanent income declines toward a minimum (subsistence) level it is hard, or impossible, for the poor to imagine an even lower consumption level. In technical terms this means that the elasticity of the marginal utility of consumption goes towards infinity, which in turn means that the poor becomes both extremely risk-averse and at the same time are not willing to save today even though it may lead to a higher income in the future (Hatlebakk, 2012). Poverty may also be explained by time inconsistency, as described in the fast growing literature on behavioral economics (see for example Banerjee and Mullainathan, 2010). Chronic poverty over generations will in that case require that lack of self-control, or other variations on time-inconsistent behavior, is transferred from one generation to another.

In the present paper we focus on credit constraints and the interaction between long-term dynamics at the household and economy level as modeled by Banerjee and Newman (1993). In the static solution households are trapped in poverty because they do not have the necessary assets to invest in new technology. Then the model describes how the distribution of assets may affect the development of the economy over time, and by that the number of households that are trapped in poverty. Banerjee and Newman show that different long run dynamics are possible, but the general conclusion is that initial distribution of assets (and thus collateral) determines the long run equilibrium of the economy. A special case of the model was presented by Ghatak and Jiang (2002), where the long run equilibrium is determined by the share of households that have too little assets to invest. The median asset will thus determine whether the economy stagnates or not. If the median is high, then few households will be credit constrained and the economy will develop.

We will test this model on data from rural India. We have collected family histories that span three generations from 800 randomly selected households in two districts of Orissa. Most of these households have lived in the same village for generations, and we investigate to what extent the dynamics of a poor household depends on the initial asset distribution of the village.

The paper adds to a growing literature on poverty dynamics that can be separated into three strands. Economists tend to use panel data for large samples but over a relatively short time-horizon, see Krishna and Shariff (2010) on India, and Dercon and Shapiro (2007) for a survey of the literature. Some economists have also studied village or state panels, see Lanjouw and Murgai (2009), Ravallion and Datt (2002), Eswaran et al. (2009), and Deaton and Dreze (2002), while others have conducted more detailed village studies, see in particular Jayaraman and Lanjouw (1999). For reviews of the literature see Baulch and Hoddinott (2000), Addison, Hulme and Kanbur (2009) and Barrett and Carter (2012). The study most similar to ours is Baulch and Davis (2008), as they also use recall of family histories, but from what we can see they do not collect information on previous generations. In contrast to the previous literature we describe, and attempt to explain, the life-trajectories over three generations for a large random sample of households. We are not aware of any previous study of this kind, and we believe the approach is useful for analysis of the deep mechanisms that may explain why some households stay poor, while others climb the ladder.

The next section discusses the underlying theoretical model and explains how we will test the model on the data. Section 3 presents the data as well as descriptive statistics on the occupational ladder today and at grandfather's time where we use a combination of assets and incomes to rank occupations. Section 4 presents the empirical analysis, while Section 5 concludes.

2. Theoretical and empirical approach

We investigate whether there are household and/or village level poverty traps in rural Orissa. The main hypothesis is that poor households are credit constrained, and thus cannot invest in modern income opportunities, whether that is by way of technology and physical capital, or human capital. There may also be a village level poverty trap, where we will apply a model by Ghatak and Jiang (2002), which is a simplified version (which allows for more specific predictions that we will test below) of Banerjee and Newman (1993). A village level poverty trap arises if there is a large share of credit constrained households in the village (and the village is not integrated with the larger economy). In that case there will be relatively few entrepreneurs and many unskilled workers, and as a result the workers will be in low demand and in equilibrium accept a low wage that is determined by their outside option as subsistence producers[1]. If the share of credit constrained (poor) households is low, on the other hand, then workers will be in high demand, wages will increase (in the Ghatak and Jiang model it will switch ones from the subsistence level to a high wage) and workers will accumulate assets and cross the barrier for investment in modern technology. In the long run equilibrium workers may thus become entrepreneurs and end up with the same high income.

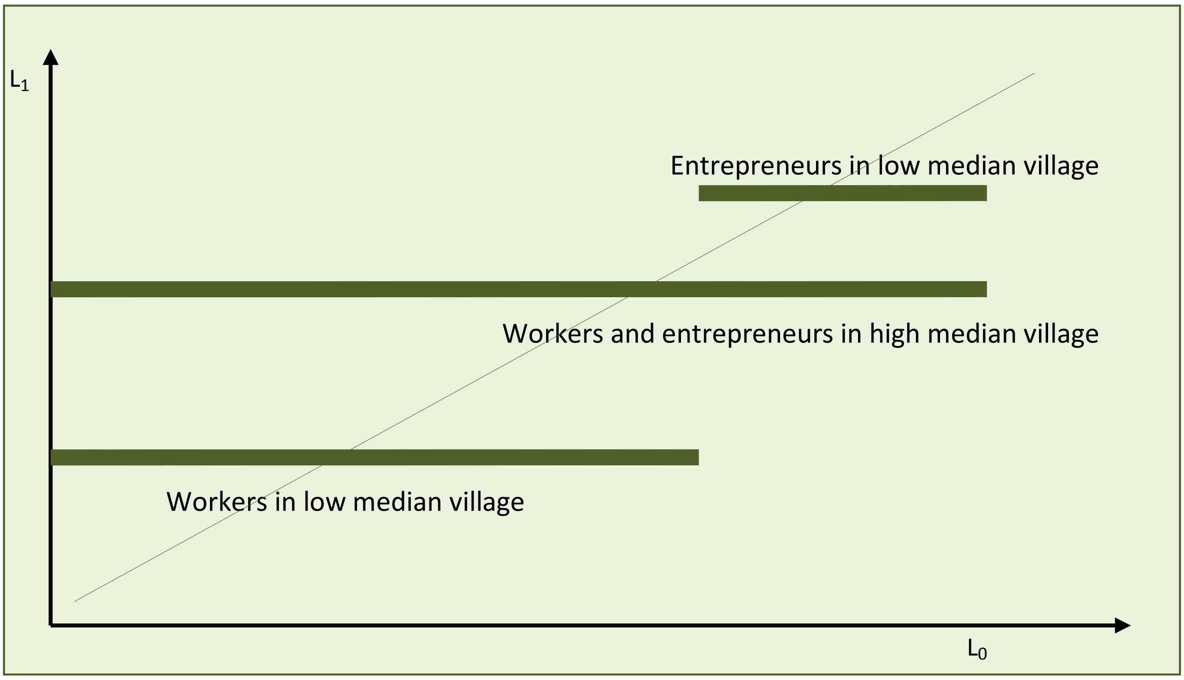
Note that it is the share of households with assets above the critical value needed for investment that, within the model, determines whether a village will stagnate or develop. We will assume that the higher is the median asset the larger is the share of households having assets above the investment threshold. This is a weak assumption, as it only means that assets are relatively evenly distributed, so that a higher median implies that at least one household crosses the threshold[2]. As a result we have that the median initial asset determines whether a village develops or stagnates.

So, within the Ghatak and Jiang model all households in a developing (high median) economy will in the long run end up with the same assets, that is, the final household asset, which we write as L_1, will not depend on the initial asset L_0. While if the median M is low, we will in the long run have two final values of L_1, a high value for the entrepreneurs that had a high value of the initial asset L_0, and a low final value of L_1 for the workers that started out with a low value of the initial asset L_0. We have illustrated these asset dynamics in Figure 1.

[1] Note that this static part of the model has parallels to, for example, Eswaran and Kotwal (1986).

[2] It is, of course, conceivable to imagine counterexamples, but in a cross section of villages it is likely that villages with a higher median will have more households above a fixed asset value.

Figure 1



If we have a cross-section of village economies, we shall expect to estimate a flat curve (as illustrated by the horizontal line in the figure) if all villages have a high median. If all villages have a low median we shall expect to estimate an increasing curve (as illustrated by the upward sloping curve in the figure). Since we do not know exactly what is a "high" and "low" median we will allow the slope (and thus also the intercept) of the curve to be a continuous function of the median:

$$L_1 = \beta_0 + \beta_1 M + \beta_2 L_0 + \beta_3 M L_0 \,. \qquad (1)$$

As an illustration assume that the median land holding M equals zero in low median villages, and equals one in high median villages, then the model reduces to $L_1 = \beta_0 + \beta_2 L_0$ in the low median villages, which would be a standard model for asset accumulation where we expect the intercept to be close to zero, and the slope to be positive, β_2 > 0. While the model for the high median villages becomes $L_1 = \beta_0 + \beta_1 + (\beta_2 + \beta_3) L_0$, where we expect a larger intercept, which implies β_1 > 0, and a horizontal slope, which implies β_3 < 0. In reality there will be more variation in M and L_1, and variation in observable (including some that we will add in the analysis) and unobservable factors, but still we keep the hypothesis that the median land holding matters, which implies the hypotheses β_1 > 0, β_2 > 0 and β_3 < 0, when we estimate (1) including the essential interaction effect $M \cdot L_0$. The median land holdings come from the descriptive statistics at the village level that will be reported in Table 7.

When we test the model on a cross-section of villages we assume that each village is a separate economy. But as the economy develops it will tend to become part of a larger regional economy, and the share of households being able to invest in the larger economy may be sufficient to develop the full region. In that case the median initial asset will have no effect on land dynamics, and we shall see that this is the finding for Cuttack district, which is close to the state capital Bubaneswar. For the remote district of Kalahandi, on the other hand, we get support for the hypotheses formulated above.

Note that a high initial median asset can be the result of previous development so that all households already are wealthy, or it can be the result of an even distribution of assets in a poor economy. In both cases we shall expect (further) development as described by the model. An alternative hypothesis would be that only the mean asset matters, and not the distribution. We will test this hypothesis in one regression by adding the mean.

The Ghatak and Jiang (2002) model also describes the corresponding occupational dynamics. In the stagnating economy households basically stay in their original occupation. There may be transfers between subsistence production and wage labor, but few transfers into the entrepreneurial class. In the developing economy, on the other hand, there will be transfers from the working class to the entrepreneurial class, and later free transfers between the two, and the subsistence class will vanish. There will still be workers, but they will have a higher wage, so that the ranking of occupations will also change. In reality there will of course be more occupations than the model describes, and the original model by Banerjee and Newman (1993) in fact has more occupations. But even with more occupations we expect to find transitions along these lines, in particular more upward transition in developing villages with a high initial median asset. And below we will also test these predictions.

The model is obviously stylized, and in the empirical analysis we will allow for other factors to affect the dynamics. The Banerjee and Newman model allows for transitions between occupations due to random events that lead to changes in assets by coincidence. We have data on events and will investigate whether events, that are more or less random, can explain some of the household dynamics. We will also investigate the importance of social identity, inspired by Akerlof (1976) as well as our own previous work in South-Asia (Hatlebakk, 2009, 2011), and initial analysis of the data that indicates that the lower class of laborers to a larger extent is made up of low caste households, scheduled tribes in the remote parts of Orissa, and some of the scheduled castes in the centrally located parts. This may indicate that social identity also plays a role in addition to the asset distribution. We will investigate not only the role of the household's own social identity, but also the caste composition at the village level.

3. Data and descriptive statistics

We collected retrospective life-histories for 800 households in Cuttack and Kalahandi districts of Orissa. The poverty level of Kalahandi is among the highest in the world, Lanjouw and Murgai (2009) report 73% poverty in the Southern districts of Orissa, with Kalahandi not necessarily being in this Southern category, but probably equally poor. Kalahandi is 600 km by road from the state capital of Bhubaneswar, while Cuttack is adjacent to Bhubaneswar, and some of the selected villages were only 30 minutes drive from the city.

We selected at random 40 households from two random villages in five randomly selected blocks in each of the two districts, in total 800 households. Households that were not found were randomly replaced (19 households, slightly more than 2%). We also had to (randomly) replace one village in Cuttack as the locals protested against any outside intervention (which would include our survey according to our field supervisors) because of a conflict regarding a power-plant. And we had to (randomly) replace a block in Kalahandi for security reasons, as it was expected to be under Maoist control. All random selection was self-weighted with probabilities according to number of households, except that we have 400 from each district, and probability weights are used to adjust for different district populations whenever we pool the data. We correct all standard errors for intra-village dependency[3].

We did one interview in each household but allowed for more than one person to participate in the interview. The interview started out with one respondent, preferably the household head, but quite often the spouse, son or another relative or household member participated. If the household head was male and at least 30 years old, then we defined him as the focal point for the family history, whether he was interviewed, or not. If the household head was a woman (or a male below 30), then we identified the husband (or father). If that person again was below 30, then we identified the father as the focal point. The last person in this sequence of logical checks would be a man of at least 30 years old that was defined as the focal point. In collecting the family history we asked detailed questions about him, his father and grandfather, allowing for the possibility that they were not alive. In addition to the family history we collected information on present day occupation of household and family members as well as a number of different assets. The questionnaire is available on request.

The family history was collected using basically one page of the questionnaire. Here we listed the names of the grandfather, father and son. Then their education, main and secondary occupation during the decade between age 30 and 40, land holding in acres at age 40, migration periods during their life, and an open ended question on events that changed the life of each of the three persons. The events were coded after the survey by the field supervisors into 56 codes, and we in turn made 16 broader categories of events. The recall problem is minimized by focusing on major events that people remember, including how much land they owned and their main occupation, we do not (in the paper) focus on the particular year of the reported events, only whether it happened at grandfather's time.

As the father and son tend to be economically active at the same time we will focus on the transition from the grandfather's economic position to the son's position. So we investigate to what extent initial endowments and events at grandfather's time explain the present economic position of the household.

[3] We use survey commands in Stata with village being the cluster, or primary sampling unit (PSU), and the districts being separate strata. In some villages there were very few households, so that most households were in the sample. Many of these have common fathers and grandfathers. The cluster option will correct for correlation between these households as well.

We classify households at grandfather's time based on his occupation and landholdings. For each occupational group we calculate the mean landholding and rank the groups accordingly. Based on the means and confidence intervals for these means we have identified four major classes that are described in Table 1. Note that we have combined some categories that are similar, even though they have significantly different landholdings, just to reduce the number of classes to four.

Table 1. Class-ranking at grandfather's time

Class	Sub-groups	N	Mean ha sub-groups	Mean ha class
1	Landlords (10+ ha)	65	30.98 (11.47-50.48)	11.57 (6.66-16.48)
	Large farmers (2-10 ha)	164	4.14 (3.84-4.43)	
2	Self-employed	75	1.49 (0.75-2.24)	1.23 (1.01-1.44)
	Medium farmers (1-2 ha)	95	1.44 (1.38-1.50)	
	Salaried work	10	1.42 (0.68-2.16)	
	Small farmers (0.5-1 ha)	84	0.74 (0.69-0.78)	
3	Marginal farmers (less than 0.5 ha)	82	0.26 (0.22-0.30)	0.25 (0.20-0.31)
	Laborers	27	0.22 (0.06-0.38)	
4	Farm-laborers	178	0.14 (0.05-0.23)	0.14 (0.05-0.23)
0	Inactive	20	1.19 (0.61-1.78)	1.19 (0.61-1.78)
		800	N=791	N=791

95%-confidence intervals (cluster corrected) in parenthesis

Grandfathers who worked as farm laborers as their main occupation when they were 30-40 years old had on average 0.14 hectares of land at age 40. While the next class of marginal farmers and non-farm laborers had on average 0.25 hectares. The third class, that we may term as the middle class of medium size farmers and the self employed and salaried employees, had on average 1.23 hectares of land. While social group one of large land owners had on average 11.57 hectares of land. Now, some households in social group two may have had higher incomes than some of the farmers in social group 1, but on average we believe this to be a useful categorization in a predominantly farming society (only 15% had the main occupation outside agriculture). However, while interpreting the results below we should keep in mind that households with large non-farm incomes are classified in social group 2.

We want to study the transitions from this traditional farming society to the present day rural society. As many as 93% of the grandfathers were born in the same village where we did the interview, so what we study here is the development over generations of households that have been living in the same village.

For the present generation we want to go beyond land in classifying and ranking the households. And we will take into consideration the occupation of all household members. So we classify each member of the household based on their main occupation. Then we add land as a criterion to split those who report farming as the main occupation into different groups based on the size of the household land. For salaried work we use income, and split the group at the 75-percentile of monthly income, which is 7000 rupees, and we do not include those who earn less than 1000 rupees in the salaried category. Factory and construction workers have on average higher daily wages than "other" non-farm workers and we separate these two group.

The self-employed is a large group, and potentially very diverse. So we have decided to split the group based on an asset index. We also use the same asset index, which now includes more than land, to rank the individual occupations that we have just described[4]. Based on this ranking we define each household's class position based on the individual with the highest rank. Then we can calculate the average value of the asset index for each class, which is reported in Table 2. This table is similar to Table 1, just that the ranking procedure is more complex.

We have information on many assets for the present day generation, including monetary values. But it is our impression that these values are not very precise, in particular at the upper end. For landholdings and the combined value of house and houseplot there are some very high values for households that otherwise do not have many assets[5]. So we have decided not to use the aggregate monetary value of assets. In stead we use principal component analysis (PCA), and use the score for the first principal component as an asset index. The self-employed individuals we separate into two groups at the value zero of the PCA-index[6]. Based on Filmer and Pritchett (2001) the PCA-index has in particular been used on DHS data, which similar to us include asset data but not income or expenditure data.

The first principal component is basically an underlying variable (which we interpret as the asset index) that is perfectly determined by the assets in the data and in such a way that it explains as much of the variation in the data as possible. So together the assets determine the index, and some of the assets will be more correlated with the index than other assets. In our case some luxury goods have a very high correlation (correlation of 0.9 for ceiling fan, followed by a number of other electrical appliances), but also major assets such as land and house characteristics are highly correlated with the index (in the range of 0.5).

[4] Here we do not use the survey commands as we consider this just a ranking of the observations in the data and not a description of the population.

[5] For rainfed land the median value per acre for landholdings of size 1-2 acres is 60 000 rupees. In such a village the land price for rainfed land will tend to vary from 40 to 80 000 rupees, which is reasonable. For irrigated land the corresponding median is in the range of 90 000 rupees, with variation in those villages from 80 to 100 000 rupees. However, in particular for irrigated land it seems to be important outliers. Among the 9 households with land value at or above 1 million rupees, 4 report a price per acre of irrigated land above 500 000 rupees. Also for house value there are 14 outliers with a house value at or above 1 million rupee. And for other assets there are again two vehicles with value above 1 million rupees. To avoid that such large values dominate the analysis, we decided to construct an asset index. There is, however, a 0.8 correlation between the monetary value of all assets and the asset index. When we look at the cases where wealthy households are classified as smallholders, their wealth consists of land and in particular the houseplot, where the value in both cases can be over-reported. So the construction of the asset index is to avoid outliers that may be due to measurement problems.

[6] Although there is strong overlap in wealth between the two groups, we find also for this group a 0.8 correlation between wealth and the asset index.

Table 2. Class-ranking at present time

Class	Sub-groups	N	Mean pca sub-groups	Mean pca class
1	High salary	68	4.71 (3.56; 5.87)	4.10 (3.41; 4.79)
	Large business	53	3.31 (2.23; 4.38)	
2	Large farmers (2+ ha)	24	1.08 (-0.19; 2.34)	0.18 (-0.32; 0.67)
	Low salary	85	0.47 (-0.43; 1.36)	
	Medium farmers (1-2 ha)	51	0.26 (-0.58; 1.10)	
	Small farmers (0.5-1 ha)	79	-0.50 (-0.91; -0.01)	
3	Marginal farmers (less than 0.5 ha)	145	-0.79 (-1.10; -0.48)	-0.92 (-1.15; -0.70)
	Factory/const. laborers	38	-1.02 (-1.35; -0.68)	
	Small business	44	-1.29 (-1.48; -1.11)	
4	Other laborers	42	-1.58 (-1.78; -1.38)	-1.50 (-1.58; -1.42)
	Farm-laborers	143	-1.48 (-1.58; -1.38)	
0	Inactive	28	-1.00 (-1.37; -0.64)	-1.00 (-1.37; -0.64)
		800	N=800	N=800

Robust to clustering 95%-confidence intervals in parenthesis

When we rank households based on the highest ranked individual occupation there is a switch in the position of farm-workers. This is because some of the farm workers live in households with people of higher rank, for example people who reported farming as the main occupation in households with some land. So marginal farm households where there are some farm-workers in the household have slightly (but not significantly) higher rank than the non-farm laborers.

As for Table 1, we make four classes based on the asset index. There is no overlap in the confidence intervals for the mean asset index for these groups, but for the sub-groups there is overlap, so the grouping is not clear-cut. The social classes have intuitive interpretations with businessmen and salaried people in social group one, farmers and low salary people in social group two, marginal farmers, factory and construction workers and petty traders in social group three, and other laborers including farm laborers at the bottom of the hierarchy.

4. Empirical analysis

This section will be organized in the following way. First we report on village level data for the 20 villages, where we include the median and mean wage today as well as the historic median and mean land holding. As we recall from Section 2 we expect the median land holding to explain the present wage level. However, 20 villages are not many, so the main test of the model will be the analysis of the household dynamics of the 800 households that will follow in separate sub-sections for land and occupational dynamics.

We start out with wages at the village level. In the theoretical model the high wage economy will also be modern, and have no subsistence production. As discussed, such a modernization may also apply to the agricultural sector. So we conduct a conservative comparison of villages by reporting the present wage of adult male agricultural laborers only. There are 558 male farm laborers from 278 households in the sample of 800 households. One of the 20 villages has only two farm laborers, which in itself indicates that the village has modernized. And the wage they are paid is also high, so even for this village it appears that we can use the wage as a measure of stagnation versus modernization. Table 3 list the number of male farm laborers and the mean and median wage among them for the 20 villages. Remember that there are 40 households interviewed in each village. We name the block, but keep the villages anonymous, so there will be two lines with the same block name.

There are approximately double number of farm laborers in Kalahandi as compared to Cuttack, which in itself indicates that Cuttack is the more modern economy. This is also rather obvious given the location near the state capital, while Kalahandi is a very remote district. Within Kalahandi there is very little variation in wages, but two villages do have slightly higher wages of 80-90 rupees, while in the other villages the wage rate is 70 rupees. In Cuttack there are two villages with low wages in the 60-70 rupees range, the same as in Kalahandi, and two villages with higher wages, in the 120 rupees range, while the normal wage in Cuttack seems to be 100 rupees. The within district variation cannot immediately be explained by studying the geographical distribution, the high wage villages do not seem to be particularly centrally located, and the low wage villages do not seem particularly remote. We do not have good price data from this survey, but when we compare prices for different livestock they are in the same range for the two districts. If we shall use this limited data to separate villages as stagnating or not, then it seems like the two villages in Baranga block in Cuttack is stagnating, and most of the villages in Kalahandi. But this is crude data and we will put more weight on the analysis of household dynamics.

We now go on to the land distribution at grandfather's time, but still at the village level. First recall that 93% of the grandfathers were born in the same village where we did the interview. There is however variation in the year of birth, so when we ask for the land owned when the grandfather was 40 years old we have observations at different points in time. We decided to ask in this manner as we expected it to be easier for the respondent to recall the land holdings of the grandfather at the height of his carrier (defined by the age of 40), than to remember the landholdings of the family at a certain point in time, let us say in 1945 (the median year when the grandfathers were 40). The latter is a long time ago, and they may even have problems in determining who was the head of the household at that time, and how the land was split between family members. But they have probably heard multiple times their grandfather telling them how much land he had when he was in charge of the family. And they probably know whether he had taken over responsibility from his own father at that age, and whether he had transferred responsibility to his sons. We find that 90% of the grandfathers were born between 1880 and 1928, meaning that the land information for this generation is for the period from 1920 to 1968. This variation is to a large extent explained by the variation in the present age of the grandson.

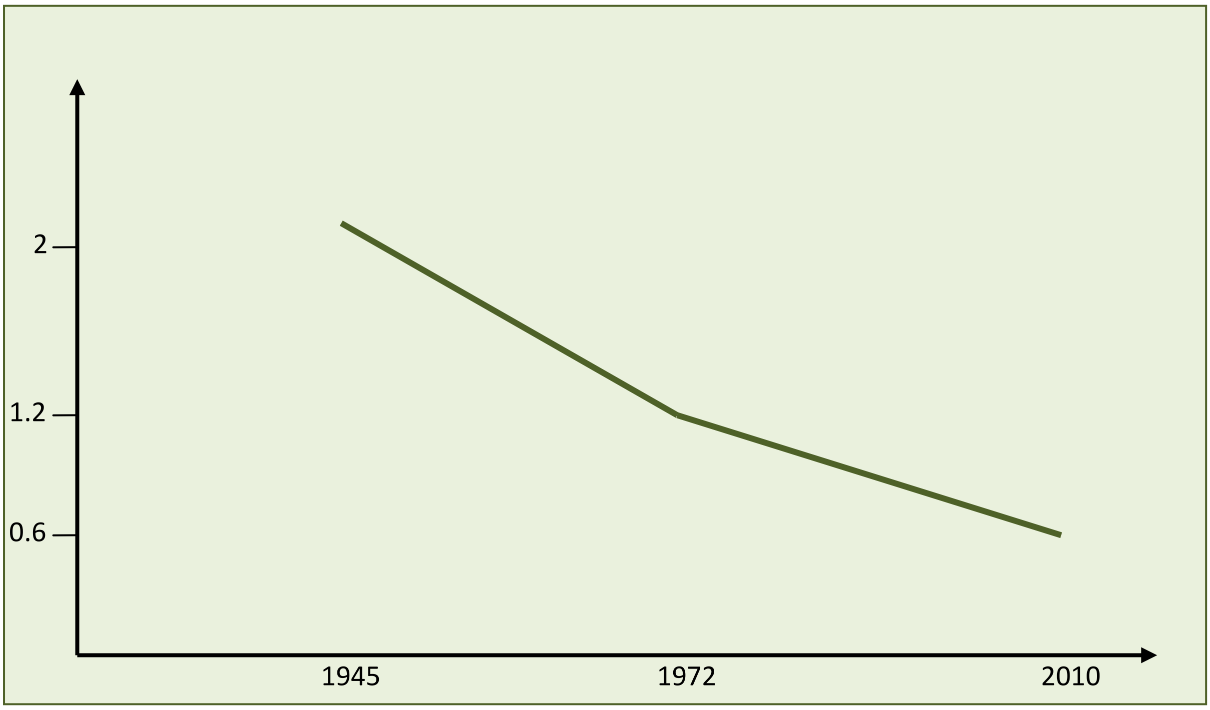
Table 3. Male farm-laborers in present generation and land distribution two generations ago

Blocks where the villages are located:	Number	Median wage in 2010	Mean wage in 2010	Median hectare in 1940	Mean hectare in 1940
Cuttack					
Baranga	47	60	61	0.46	0.94
Baranga	10	70	75	0.40	0.81
Badamba	26	100	96	1.09	1.41
Badamba	14	100	94	0.51	0.90
Mahanga	18	120	115	0.61	0.93
Mahanga	12	90	93	0.87	1.74
Niala	11	100	100	0.75	1.81
Niala	9	100	102	1.14	1.74
Nischintakoili	44	100	98	0.29	0.33
Nischintakoili	2	120	120	0.50	1.00
Kalahandi					
Dharamgarh	34	78	80	0	3.83
Dharamgarh	28	70	67	3.12	5.24
Jayapatna	52	70	70	1.94	13.56
Jayapatna	40	70	70	2.19	12.91
Junagarh	60	70	68	0	3.23
Junagarh	32	70	71	2.66	6.64
Madanpur Rampur	39	80	89	5.51	10.03
Madanpur Rampur	10	70	70	1.05	2.34
Narla	40	70	71	2.91	8.57
Narla	30	70	75	0.90	4.04

We now go on to the land distribution at grandfather's time, but still at the village level. First recall that 93% of the grandfathers were born in the same village where we did the interview. There is however variation in the year of birth, so when we ask for the land owned when the grandfather was 40 years old we have observations at different points in time. We decided to ask in this manner as we expected it to be easier for the respondent to recall the land holdings of the grandfather at the height of his carrier (defined by the age of 40), than to remember the landholdings of the family at a certain point in time, let us say in 1945 (the median year when the grandfathers were 40). The latter is a long time ago, and they may even have problems in determining who was the head of the household at that time, and how the land was split between family members. But they have probably heard multiple times their grandfather telling them how much land he had when he was in charge of the family. And they probably know whether he had taken over responsibility from his own father at that age, and whether he had transferred responsibility to his sons. We find that 90% of the grandfathers were born between 1880 and 1928, meaning that the land information for this generation is for the period from 1920 to 1968. This variation is to a large extent explained by the variation in the present age of the grandson.

When we analyze individual household dynamics below we focus on historic land holdings as a predictor for present economic status. But we will have to adjust for the number of years that have past since the grandfather was 40. We may add number of years as an explanatory variable, but a more direct way will be to adjust the measure of land at grandfather's time, using the information we have on land in later generations. Now, the loss of land is for most households larger from grandfather's time to father's time than from father's time till today. So land is basically a downward-sloping convex function of time. To simplify we depict this by two linear segments between the three observations. This allows for flexibility at the individual level, in stead of forcing all observations into a joint function. In addition we assume that the linear segment continues back in time from the earliest point for which we have information. By this approach we get an estimate for the amount of land of a household at any point in time. The approach is illustrated in Figure 2, where a household's land at three different points in time is represented by the medians for both variables for the three generations.

Figure 2.



Now the figure only depicts the dynamics of one family. Wealthier families will in particular tend to have a steeper first part. If we now want the land distribution prior to 1945, and since this is the median, we have 50% of the observations in that range, we would have to stipulate the landholdings of the other 50% by extrapolating the left part of the line. If we want the distribution at later years stipulate backwards in time will necessary for fewer families. For families with no information on the grandfather we would have to stipulate based on the second line. For any particular year, we will however need to apply the linear trend assumption to identify the landholding. Using the example from the figure again, this particular household will, for example, have approximately 1.6 acres in 1960. In the analysis we want to pick a year that has the largest concentration of grandfathers, so the natural choice is the median of 1945. We have that 50% of the grandfather's will be in the range from 1935 to 1953, so a good concentration around the median. But since stipulation to the left may be more erroneous due to the likely convex underlying function we will rather use 1940 (which is also a more common year). So we stipulate the household landholdings in 1940 based on linear extrapolation, and investigate how that influences landholdings 70 years later in 2010.

The median and mean landholdings in 1940 are reported in Table 3. The land distribution is most equal in Cuttack as measured by the low means as compared to the median. The median to mean ratio goes from 0.41 in a village in the block of Niala to 0.89 in a village in Nischintakoili, which implies that the latter had the most equal land distribution. The mean hectare goes from 0.33 hectares in the same village to 1.81 in the village with the unequal land distribution. But keep in mind that land values were probably higher in Cuttack, and also that people to a larger extent relied on non-farm incomes, so income inequality may have been higher.

In Kalahandi there are villages where the median land holding were zero, so by that the median to mean ratio is zero as well. This is the case for a village in Dharamgarh block and another village in Junagarh block. The village with least inequality by this measure is the second village in Dharamgarh block where the median to mean ratio is 0.6. In Kalahandi there appears to be more variation between villages, so the village with the most equal land distribution is not the one with lowest mean. So here we may expect that the mean (as a measure of development) and the median (as a measure of inequality when we control for the mean) may have separate explanatory power. The villages with highest mean are the two in Jayapatna block and one in Madanpur Rampur, with means from 10 hectares to 13 hectares.

Now the prediction from the model is that if the historical median land holding is below a critical threshold then the local economy will end up in a low wage equilibrium. Alternatively one may imagine that the mean land holding is more essential. So a crude test of competing theories can be to regress the median wage in Table 3 against the historical median and mean land holding, potentially separately for the two districts. Now note that the village will be the unit of observation, so we will have only 10 observations per district. As we know, we do have the underlying household observations, and below we will utilize the full data to analyze household dynamics. But for this particular hypothesis regarding the implication for the village level equilibrium wages, we would inflate the data by using households as the unit of observation.

If we look at the correlation between the variables in Table 3 we find that if we pool the two districts then the mean land holdings would appear to have a long term negative effect as farm wages are lower in Kalahandi and land holdings were larger. But the higher wages in Cuttack are probably explained by factors outside the farm sector, so we have to separate the districts. If we do so then there is no significant effect of the mean land holding on wages, with correlations being 0.14 or lower, and with no significant parameter in a simple OLS. For the median land holding there are higher correlations, and in particular so for Kalahandi, where the positive correlation between the median land holding in 1940 and the mean farm wage in 2010 is 0.41. But as the table shows there is large variation (and a small sample size) so in an OLS the parameter is not significant.

But if we eyeball Table 3 and focus on the villages in Kalahandi we see that we have the highest wage of 89 rupees in the village (in Madanpur Rampur) where the median was the highest at 5.51 hectares. So this finding is in support of a model where this particular village has modernized, while other villages with low wages and a low historic median land holding have stagnated, such as the village in Junagarh where the average wage is 68 rupees and the median land holding was zero. But let us now go on to the proper statistical test of the theoretical model, where we apply the full data of 800 households to test the model of household dynamics.

4.1 Household asset dynamics

We focus on land only, and not the asset index that we used in the occupational ranking above, since we have information on other assets only for the present generation. As described in Section 2 the

Ghatak and Jiang model predicts that if the median historic land holding is high then the village will develop and workers and entrepreneurs will end up with the same assets. If the median is low then the village will stagnate and there will be two levels of assets in the long run equilibrium. We estimate a regression model where present land (in 2010) is a function of the historic land holding (in 1940). If the villages have in fact ended up in the long run equilibrium, and otherwise are similar, then the present land will be independent of historic land in the high median villages, and we will get two horizontal lines in the low median villages, as illustrated above in Figure 1. If this is the data generating process, then an OLS for the high wage villages will give a horizontal regression line, while an OLS for the low wage villages will give an increasing line as depicted in the figure and discussed in Section 2. If villages are not in the long run equilibrium, and if there are other factors that determine household asset dynamics, then we shall expect data point scattered in the diagram. But if the model has explanatory power, we shall still be able to identify an underlying data generating process of this kind.

Before we present the findings, there is one more issue that may be of concern. Both historic and present land holdings include many zeros. Zeros on the independent variable is normally not considered an econometric problem, it only means many observations in the intercept with the vertical axis. But if we estimate a normal OLS we would automatically assume that households with initial values near zero would behave in a similar manner to landless households. This is a strong assumption as households with some land may consider to accumulate land, while landless households may not. In some of the regressions we thus add a dummy for landlessness, and also an interaction of the dummy with the median. A dummy for landlessness means that we allow those observations (that will all be on the vertical axis) to not necessarily be positioned in the intercept of the regression line for the households with positive land. And allowing for the interaction of the dummy with the median, we may in principle end up with multiple points on the vertical axis.

Many zeros for the dependent variable raise the issue of different accumulation processes for households ending up as either landless or with land. Note first that a tobit is not the right model as we cannot conceive of a latent variable with negative values, that are reported as zero. But since the processes that lead to landlessness or not may differ, we have also run regressions where the present land is measured as zero or one, where one represents all positive land values, and the regression model in (1) will turn into a linear probability model. Again we expect, and find, the same signs for the parameters, that is, the probability of having land (versus being landless) to depend more on initial land in villages with a low median[7]. On the aggregate we find that landlessness declined in Kalahandi from 32% to 29%, while it increased in Cuttack from 17% to 26%. The increase in Cuttack indicates that households left the farming sector, and with Cuttack villages being more integrated with the urban economy we are, as discussed above, less convinced that our test of the model will apply there.

Note that the land dynamics are actually not increases in land, on average the present land holding is much smaller than the historic land holding. But that does not change the hypothesis, we just have to keep in mind that households that do well, have a smaller decline in land. On average we find that land declined from 3.4 hectares in 1940 to 0.5 in 2010.

We will also, in some regressions, add control variables that will shift the intercept in Figure 1, but not the slope, so for a certain initial land holding, some groups of households may end up with more land

[7] These regressions with a dichotomous dependent variable were included in earlier versions of this paper, available on request, but are not included here since the findings are basically the same.

today. We add own caste[8], as well as the caste composition of the village, and in a separate regression, indicators of early economic shocks as reported by the households. The latter may be endogenous, so those OLS regressions should be considered as tentative indications that these shocks do not matter much. By including village caste composition in the regression we can separate the effect of own caste on change in land from the effect of the composition of the village[9].

Finally note that we add the mean land holding in one regression. This is because one may imagine that the mean land holding is so high that all households, independently of the distribution, can become entrepreneurs. But as we shall see the mean does not change the findings for Kalahandi, where the median is the essential variable. In Cuttack on the other hand, where we do not find support for the theoretical model, the mean is more important than the median. The variables are presented in Table 4 and the regressions in Table 5 and 6.

[8] In reality the main difference turn up for the so called upper castes, so we only add a dummy for this "general" caste.

[9] We may potentially add number of sons of the grandfather, but this does not change the findings. The variable is significant, in non-reported regressions, but only for Kalahandi, which again indicates that Cuttack is a modern economy where they get fewer sons, and also where sons can find non farm occupations and thus may not take their share of the land.

Table 4. Dependent and independent variables

	Kalahandi (N=398)				Cuttack (N=399)			
VARIABLES	mean	st.dev.	min	max	mean	st,dev.	min	max
Dependent:								
hectare-present	0.670	0.869	0	6.475	0.342	0.485	0	4.047
Landholder-present	0.709	0.455	0	1	0.739	0.440	0	1
Independent:								
Initially landless	0.324	0.469	0	1	0.165	0.372	0	1
hectare-initial	5.967	12.110	0	88.239	1.162	1.817	0	17.851
median-village-hectare	2.025	1.586	0	5.506	0.663	0.275	0.293	1.138
mean-village-hectare	7.019	3.841	2.336	13.563	1.162	0.468	0.329	1.807
interaction	15.401	40.974	0	485.877	0.879	1.543	0	13.366
interaction-w/zero	0.421	0.980	0	5.506	0.089	0.223	0	1.138
general caste	0.055	0.229	0	1	0.113	0.317	0	1
%-general caste	0.058	0.069	0	0.175	0.112	0.114	0	0.325
early-nat-disaster	0.048	0.213	0	1	0.100	0.301	0	1
early-health-prob	0.126	0.332	0	1	0.035	0.184	0	1

In Kalahandi the mean landholding has declined from 5.97 hectares in 1940 to 0.67 hectares in 2010, while the proportion of landless households has been relatively constant with 32% landless in 1940 and (1 - 0.709 =) 29% landless in 2010. In Cuttack the mean landholding has declined from 1.16 hectares in 1940 to 0.34 hectares in 2010, while the proportion of landless households has increased from 17% landless in 1940 to (1 - 0.739 =) 26% landless in 2010. The latter is probably explained by a switch to non-farm occupations, but also by the smaller land holdings in general in Cuttack. This difference also shows up in the aggregate, the median initial land holding (in fact the mean of the calculated medians) was 2 hectares in Kalahandi and 0.66 hectares in Cuttack. The corresponding means were 7 and 1.16 hectares[10]. The values of the interaction effects have no immediate interpretation.

Regarding caste we have 5.5% so called general or higher castes in Kalahandi and 11.3% in Cuttack[11]. And we see that 4.8% of the grandfathers in Kalahandi, and 10% in Cuttack are reported to have experienced a natural disaster ("that changed his life"), and 12.6% grandfathers in Kalahandi and only 3.5% in Cuttack experienced a health problem of similar importance. These differences in shocks may reflect that Cuttack is a richer district, both at the individual level, but maybe also at the aggregate (or public level) level in terms of access to health services. While on the other hand, Cuttack is in particular more prone to cyclones and floods. We now report on the regression analysis.

[10] We have omitted one outlier from the analysis that had more than 400 hectares initial land. The second largest was less than 100 hectares. Including the single outlier drastically reduces the importance of initial land. The observation is included in calculation of the mean and median though as the household may have had an effect on the development of the village. The latter explains the higher mean of the means in Kalahandi.

[11] The mean of the means for this variable is also slightly different as the proportions are calculated for all households while three households are dropped, one outlier and two because we do not have the necessary information to calculate the initial land holding.

Table 5. Land in 2010 as a function of land in 1940, Kalahandi

VARIABLES	(1) hectare	(2) hectare	(3) hectare	(4) hectare	(5) hectare	(6) hectare	(7) hectare	(8) hectare
1=zero hectinit		-0.498***	-0.589***		-0.503***			
		(0.063)	(0.090)		(0.091)			
hectare-initial	0.031**	0.024**	0.025**	0.074***	0.062***	0.065***	0.065***	0.066***
	(0.009)	(0.009)	(0.009)	(0.013)	(0.011)	(0.013)	(0.012)	(0.013)
median-village			-0.091**	0.046**	-0.012	0.055**	0.042	0.056**
			(0.033)	(0.016)	(0.020)	(0.021)	(0.027)	(0.021)
mean-village			-0.001					
			(0.017)					
interaction				-0.015***	-0.013***	-0.014***	-0.014***	-0.014***
				(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
interaction -w/zero					0.025			
					(0.022)			
general caste						0.665*	0.706**	0.658*
						(0.314)	(0.299)	(0.317)
%-general caste							-0.609	
							(0.694)	
early-nat-disaster								-0.099
								(0.186)
early-health-prob								0.035
								(0.112)
Constant	0.487***	0.687***	0.903***	0.372***	0.676***	0.343***	0.401***	0.341***
	(0.047)	(0.056)	(0.096)	(0.068)	(0.091)	(0.077)	(0.113)	(0.082)
Observations	398	398	398	398	398	398	398	398
R-squared	0.182	0.246	0.271	0.302	0.351	0.328	0.329	0.329

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

First note that historical landholdings predict present land holding, as expected. In the second regression we note that the historically landless (with L_0 = 0 which is noted as 1 = zero hectinit in the table) have 0.489 less hectares of land than the intercept for those who had land in 1940. So there is a different process of land accumulation for those who were completely landless. But still there is an upward sloping curve also for those who had land. In the third regression we add the median village land holding. We also add the mean just to see whether the median is the essential variable, and it is. Now this regression only make sense if the interaction effect is not significant, but it is, as we see in regression four, which is the exact representation of equation (1). And we see here that all hypotheses are confirmed. The intercept is higher and the slope less steep in villages with a large median land holding. So we have support for the Ghatak and Jiang model. Present land value depends less on initial land the higher is the median land holding of the village. So in those villages with a more equal land distribution it appears that everyone gets a chance as the economy develops.

The fifth regression generalizes the model to take into account the different accumulation process among the landless. The findings are basically the same, with a less steep function the higher is the median. The intercept does no longer depend significantly on the median, but that is probably because the intercept is higher when the landless are left out. The three last columns add control variables to the main model, and we see that the parameters for the main variables are robust, again with the exception for the shift in the intercept that depends on the median as this parameter is not significant in the seventh regression. But we see that the size is in the same range, so the caste composition variable appears to add some noise. Regarding the control variables themselves, we see that early shocks appear not to matter. The same is the case for caste composition of the village. But the so called general, or higher, castes seem to have a higher land holding today than households from other groups that started out with the same initial land holding.

Table 6. Land in 2010 as a function of land in 1940, Cuttack

VARIABLES	(1) hectare	(2) hectare	(3) hectare	(4) hectare	(5) hectare	(6) hectare	(7) hectare	(8) hectare
1=zero hectinit		-0.192***	-0.193***		-0.220*			
		(0.058)	(0.056)		(0.115)			
hectare-initial	0.105***	0.093***	0.089***	0.085*	0.047	0.060	0.055	0.061
	(0.013)	(0.014)	(0.015)	(0.040)	(0.038)	(0.043)	(0.040)	(0.041)
median-village			-0.279	0.015	-0.067	0.004	-0.008	0.009
			(0.215)	(0.120)	(0.144)	(0.108)	(0.104)	(0.105)
mean-village			0.200*					
			(0.103)					
interaction				0.024	0.059	0.042	0.048	0.041
				(0.053)	(0.045)	(0.058)	(0.055)	(0.056)
interaction-w/zero					0.029			
					(0.149)			
general caste						0.187**	0.165	0.178*
						(0.074)	(0.094)	(0.081)
%-general caste							0.176	
							(0.283)	
early-nat-disaster								-0.027
								(0.043)
early-health-prob								0.142
								(0.162)
Constant	0.220***	0.265***	0.223**	0.212**	0.313**	0.211**	0.202**	0.207**
	(0.033)	(0.040)	(0.075)	(0.089)	(0.106)	(0.080)	(0.074)	(0.081)
Observations	399	399	399	399	399	399	399	399
R-squared	0.153	0.173	0.184	0.154	0.175	0.167	0.169	0.171

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In Cuttack, on the other hand, we find no support for the Ghatak and Jiang model, which is in support of our discussion of the wages in Table 3. In Cuttack it is only the households own land that matters. If you were landless then you tend to stay landless, and if you had land that will determine how much land you have today. This is the same in Kalahandi, but there the village level development also matters. In Cuttack it also matters, but not via land distribution, as measured by the median, but only via the initial mean land holding, which is significant in the third regression.

Let us now summarize the asset findings. We find a significant interaction effect only for Kalahandi, indicating that in remote villages with many asset poor households the asset accumulation of each household will depend on their initial asset holding. This is not the case in Cuttack, probably because the poor have easier access to a larger labor market and thus can accumulate income independently of local labor market conditions. Note that we have a double support for the theoretical model. We expected the model to have explanatory power in Kalahandi, but not in Cuttack, and the findings support this conjecture. In Cuttack there is a direct effect of the mean initial asset level, but this applies to all households. While in the remote villages of Kalahandi the poor stay poor and in particularly so if they live in a village with many poor households. So we have identified a village level poverty trap that goes beyond the household level poverty trap.

4.2 Household occupational dynamics

The theoretical model also has predictions for occupational dynamics, and we test these predictions as well. On the trajectory to the high salary equilibrium we shall expect upward movements on the occupational ladder. While on the low salary trajectory there may be transitions between occupations at the same level, for example from farm labor into low paid non-farm labor, but we shall expect few upward transitions. We thus combine Table 1 and 2 into a transition table, which we thereafter categorize into types of transitions, and finally separate according to districts and median initial landholding. We will also separate the table according to caste identity as caste seems to have separate explanatory power, and we will finally check whether early events may explain some of the variation.

Recall that the sample is representative for the present generation, and not the grandfather's generation. At grandfather's time there were households that no longer exist, and some of the present day households are the offspring of the same household at grandfather's time. As a result we base all percentages on the present generation, so the transition table shows column-sums, in stead of row-sums which would be the standard in a transition matrix. Note that 46 households are inactive in at least one of the two periods, and are not included in this table.

Table 7. Transition from grandfather's time to present generation

Class grandfather	Class present generation			
	1	2	3	4
1	28 (23.5)	119 (51.3)	47 (21.2)	29 (16.0)
2	62 (52.1)	62 (26.7)	90 (40.5)	38 (21.0)
3	20 (16.8)	19 (8.2)	43 (19.4)	24 (13.3)
4	9 (7.6)	32 (13.8)	42 (18.9)	90 (49.7)
N=754	119 (100)	232 (100)	222 (100)	181 (100)

Percentages in parentheses

Here we can immediately see that 50% of the lowest class today (farm and other laborers) had a grandfather who was also a farm-laborer. So 50% of the present poor are in a poverty trap that has lasted for generations. But there is obviously another side to this coin, the 90 poverty trapped households are matched by another 83 households (the left poverty group below) with a grandfather who was a farm labor, but where the present generation is in a higher ranked social group.

We now reduce the 16 cell transition table into six transition categories that are summarized in Table 8. Note that the numbers in Table 8 are derived directly from Table 7. The into-poverty group is for example the sum of the three first entries in the last column of Table 7 (29+38+24=91).

Table 8. Transition categories

Transition	Table 3	N	%
Poverty trap	Always in 4	90	12%
Into poverty	1-3 into 4	91	12%
Left poverty	4 into 1-3	83	11%
Stagnated	Always in 1, 2 or 3	133	18%
Improved	3 to 2,1 or 2 to 1	101	13%
Declined	1 to 2, 3 or 2 to 3	256	34%
		754	100%

So while there are 50% of the poor being in a poverty trap, there are equally many households that have entered into poverty, and in fact there are about as many households that have left poverty, indicating that the social group four of farm and other low paid workers have been relatively stable over the generations in the aggregate, but with transition in and out of poverty. With a 50% split the reader may focus on the large share of the poor that is trapped in poverty, or the large transition in and out of poverty.

The largest group, though, is the 34% that are sliding down the social ladder. Within this group the largest sub-group is the 119 households that started out in social group one (farmers with at least two hectares land) and ended up in social group two, which is dominated by low salaried households and farmers with less than two hectares land. This is a general trend we see in the data, household sell or split land between sons, and they transfer into relatively low paid non-farm jobs. However, we have to keep in mind that these ladders are relative, there has been economic growth in Orissa over the two generations, so a low salaried employee, or a small farmer, today may have a living standard that is comparable to a large farmer 50 years ago[12]. We now split Table 8 between districts, and according to the median land holding. Note that there are slightly more inactive households in the high median Cuttack villages, which explains the lower sample size there.

[12] The mean reported birth year of the grandfathers in the sample is 1905, while the mean birth year for their grandsons is 1961. Now we believe that the respondents have overstated the age difference, so the average gap between the generations is probably well below 28 years, probably closer to 20.

Table 9. Transitions by district and median landholding

Transition	Kalahandi		Cuttack	
	Low median	High median	Low median	High median
Poverty trap	40 (**20.8**)	20 (10.6)	17 (8.9)	13 (7.1)
Into poverty	23 (12.0)	35 (18.5)	14 (7.3)	19 (10.4)
Left poverty	26 (13.5)	20 (10.6)	30 (**15.7**)	7 (3.9)
Stagnated	32 (**16.7**)	10 (5.3)	47 (24.6)	44 (24.2)
Improved	11 (**5.7**)	3 (1.6)	52 (27.2)	35 (19.2)
Declined	60 (**31.3**)	101 (53.4)	31 (**16.2**)	64 (35.2)
N=754	192 (100)	189 (100)	191 (100)	182 (100)

Percentages in parentheses.

Bold means significantly different from high median at the 90%-level with clustered standard errors.

We get the expected findings. For the low median villages in Kalahandi, where we shall expect to find stagnation, we do in fact find stagnation. The poverty trapped, stagnated and declining households are overrepresented. The only exception is the few (11 as compared to 3 households in the high median villages) that have climbed the ladder.

In Cuttack, on the other hand, there are few poverty trapped households, few households that have fallen into poverty, and for these categories no difference between low and high median villages. There is, however, a difference for the left poverty group. In Cuttack many more households have left poverty in the low median villages. This actually corresponds with the asset dynamics findings, as we did not find support for the low wage path in Cuttack. So if all Cuttack is in fact on the trajectory to the high wage equilibrium we shall expect to find more households leaving poverty in villages that started out with a low median land holding. The higher number of declining households in the high median villages probably reflects a similar reversion to the mean as the wealthiest employers now have to pay the workers better, as explained by the model, when all in the long run and up with similar incomes.

4.3 Caste and occupational transitions

As we have found caste to affect asset dynamics, we investigate whether caste plays a role also for occupational transitions, and split Table 8 according to the social identity of the households, as shown in Table 10. As mentioned we use the official categorization of people into scheduled caste (SC), scheduled tribe (ST), other backward castes (OBC) and others who are named as general caste. Now the classification of a particular group into these broad caste categories may change over time. We use however the present categorization as reported by the households themselves, and thus assume that low ranked groups were low ranked also two generations ago. It is beyond the scope of the paper to reclassify castes at grandfather's time.

Table 10. Transitions and social identity

	Caste			
Transition	**SC**	**ST**	**OBC**	**General**
Poverty trap	26 (15.8)	20 (**20.6**)	42 (9.5)	2 (2.8)
Into poverty	22 (12.9)	10 (10.3)	56 (12.9)	3 (**4.7**)
Left poverty	37 (**22.1**)	4 (4.1)	42 (9.6)	0 (**0.0**)
Stagnated	34 (20.3)	12 (12.4)	65 (16.0)	22 (**34.2**)
Improved	16 (**9.7**)	0 (**0.0**)	65 (16.0)	20 (**32.0**)
Declined	33 (**19.3**)	51 (**52.6**)	155 (36.0)	17 (26.3)
N=754	168 (100)	97 (100)	425 (100)	64 (100)

Percentages in parenthesis are probability weighted. Bold means significantly different from OBC

Maybe surprisingly we find that the schedule caste category is overrepresented only in the left-poverty category, while the scheduled tribes are overrepresented in the poverty-trap category as well as in the decline category. As we may expect, the general category is overrepresented in the stagnated and improve categories, which both imply that they are well off today.

4.4 Early events and occupational transitions

Above we found no effect of early events on accumulation of land, but there may still be an effect on occupational transitions. That is, there may be particular events at grandfather's time that explain the transitions. For 492 of the 800 households no event is reported for the grandfathers. The remaining 308 households report 540 events, so 1.75 events per grandfather. Now, the low number of households reporting on events at grandfather's time is not explained by lack of a willingness to respond, for events at father's time as many as 725 households report in total 1872 events, so almost all households and 2.6 events per father. So probably they just did not know the important events during grandfather's time. We will still use the data below, assuming that the events that are reported were essential for the development of the household. Among the 540 events, the most frequent are land sales (85), family separated (81), health problems (78), natural disaster (74), and started economic activity (60).

That the family separated is, in our mind, not important as that will happen at a certain point in time for all households, and some households decided to report this, others not. Furthermore, land sales and upstart of an economic activity (this includes starting a business, or in a new job, or labor migration) are in our mind important descriptions of the transition, rather than events that nay explain the transitions. This contrasts with natural disasters and health problems that are more likely to be random, and are thus events that may explain rather than describe the transitions. In Table 11 we split Table 8 according to whether the households report one or more of these events (except family separated) at grandfather's time. Note that some households are represented in more than one column as 30 households report two of these events, and one household report three of them.

Table 11. Transitions and events at grandfather's time

Transition	Early events				
	No event	Natural disaster	Health shock	Land sale	Economic activity
Poverty trap	74 (13.2)	6 (10.6)	3 (5.1)	3 (4.0)	4 (7.5)
Into poverty	49 (8.8)	9 (15.3)	14 (**22.7**)	15 (**21.2**)	15 (**28.8**)
Left poverty	67 (12.1)	3 (5.2)	6 (10.4)	2 (2.9)	7 (14.3)
Stagnated	93 (17.4)	9 (16.0)	12 (20.1)	17 (25.4)	8 (16.5)
Improved	77 (14.7)	10 (18.0)	3 (5.3)	5 (7.7)	10 (21.1)
Declined	188 (33.8)	21 (35.0)	22 (36.3)	28 (38.9)	6 (11.9)
N=754	548 (100)	58 (100)	60 (100)	70 (100)	50 (100)

Percentages in parenthesis are probability weighted. Bold means significantly different from no event

There is one robust finding here, those who have transferred into poverty report many more events than any other group. They basically report all types of events, although natural disasters are not significantly over-represented. One may believe that the poor just had more time available, or more respect for the enumerators, but the poverty trapped households do not report many events. So it is likely to be a real phenomenon. So the households that have declined into poverty had to a larger extent negative health shocks at the grandfather's time. And they also sold land and started a new economic activity to a larger extent than other households during the grandfather's generation. It is tempting to argue that the logical sequence is that the health problems implied that they had to sell land, and find another occupation. But in most cases only one of these events are reported, and in case more than one event is reported the health problems happen later in life.

So we rather conclude that sale of land is an indicator that the household is not doing so well, and health problems may tip the household towards a declining trajectory in the following generations. It also appears that starting new economic activities are indications of a downward sloping trend, which in turns indicate that the grandfather's were forced into new occupations. Or it may be the case that it is risky to start a new economic activity, but in that case we should expect this activity to be overrepresented also in the improve category. It appears overrepresented but there are only ten observations in this cell and thus no significant difference can be found. But, there is some indication that events at grandfather's time matter for the occupational dynamics

5. Conclusions

We describe poverty dynamics over three generations and investigate, using a model by Ghatak and Jiang (2002), which in turn is based on Banerjee and Newman (1993), whether early village land distribution determines the present level of economic development, and we find support for this conclusion in the remote district of Kalahandi. Villages with an unequal land distribution, and thus a low median land holding, are more likely stagnated. Within the model, and supported by the data, this means that the land dynamics of a household depends less on the household's initial land the higher is the initial median land holding of the village. In other words, the household is less likely to stay poor if the neighbors are non-poor. The underlying mechanism being that the poor will benefit from the relative lack of labor and the investments made by the non-poor, so that laborers end up with similar incomes to the entrepreneurs in the long run.

When it comes to the more detailed individual household dynamics, we find that half of the households that base their income on low paid unskilled labor today had a grandfather who was himself an unskilled farm laborer. These households are in a poverty trap that is explained by their marginal initial assets. However, we also have poor farm laborers at grandfather's time who now have grandsons that have been able to climb the economic ladder and work as factory or construction workers, or in low-salaried jobs. And we have grandfathers who were small and marginal farmers, and now have grandsons with barely any land left who work as farm laborers. We find that in particular the scheduled tribes are over-represented among the poverty trapped households, while the scheduled castes are overrepresented among the households that have left poverty. This may reflect that India has given priority to Dalits in social programs including affirmative action, while scheduled tribes, who normally live more concentrated in remote villages, have not benefitted as much, probably because it is harder to find non-agricultural employment in remote areas, which would otherwise allow unskilled labor to climb the ladder. This lack of economic progress may in turn explain the increase in Maoist activities in these areas.

The findings indicate that programs targeting the poor should be focused on remote schedule-tribe villages where there is widespread poverty rather than semi-urban areas where the poor are in a minority. Even though the number of beneficiaries may be the same, and the potential for development may seem more promising in more developed places, the analysis indicates that a large push in the most remote locations has the potential of lifting villages out of poverty traps, while poor households in more developed places can expect to benefit from modernization independently of public programs.

References

Addison, T., Hulme, D. and Kanbur, R. (2009). *Poverty Dynamics: Interdisciplinary Perspectives*. Oxford University Press.

Akerlof, G. (1976). "The Economics of Caste and of the Rat Race and other Woeful Tales." Quarterly Journal of Economics. 90: 599-617.

Banerjee, A. and Mullainathan, S. (2010). *The Shape of Temptation: Implications for the Economic Lives of the Poor*. NBER Working Paper No. 15973.

Banerjee, A.V. and Newman, A.F. (1993). "Occupational Choice and the Process of Development". *Journal of Political Economy*. **101**(2): 274-298.

Barrett, C.B. and Carter, M.R. (2012). "The Economics of Poverty Traps and Persistent Poverty: Policy and Empirical Implications". *Journal of Development Studies*. Forthcoming.

Baulch, B. and Davis, P. (2008). "Poverty Dynamics and Life Trajectories in Rural Bangladesh". *International Journal of Multiple Research Approaches*. **2**(2): 176–190.

Baulch, B and Hoddinott, J. (2000). "Economic mobility and poverty dynamics in developing countries". *Journal of Development Studies.* **36**(6): 1-45.

Deaton, A. and Dreze, J. (2002). "Poverty and Inequality in India. A Re-Examination". *Economic and Political Weekly*. **37**(36): 3729-3748.

Dercon, S. and Shapiro, J.S. (2007). "Moving on, Staying Behind, Getting Lost: Lessons on Poverty Mobility from Longitudinal Data", in Narayan D. and Petesch, P. (eds.) *Moving out of Poverty. Volume I. Cross-Disciplinary Perspectives on Mobility*. Palgrave Macmillian and the World Bank.

Eswaran, M. and Kotwal, A. (1986). "Access to Capital and Agrarian Production Organisation". *Economic Journal*. **96**(382): 482-498.

Eswaran, M., Kotwal, A., Ramaswami, B and Wadhwa, W. (2009). "Sectoral Labour Flows and Agricultural Wages in India, 1983-2004: Has Growth Trickled Down?" *Economic and Political Weekly*. **44**(2): 46-55.

Filmer, D. and Pritchett, L. (2001). "Estimating wealth effects without expenditure data, or tears: An application to educational enrollments in states of India". *Demography*. **38**(1): 115–132.

Ghatak, M. and Jiang, N. (2002). "A simple model of inequality, occupational choice, and development". *Journal of Development Economics*. **69**(1): 205– 226.

Hatlebakk, M. (2009). "Capacity-constrained Collusive Price Discrimination in the Informal Rural Credit Markets of Nepal". *Review of Development Economics.***13**(1): 70-86.

Hatlebakk, M. (2011). "Triadic Power Relations in Rural Nepal". *Journal of Development Studies*. **47**(11):1739-1756.

Hatlebakk, M. (2012). *Myopic preferences or subsistence income? Why do rickshaw cyclists rent the cycle?* CMI-WP: 2012: 1.

Jayaraman, R. and Lanjouw, P. (1999). "The Evolution of Poverty and Inequality in Indian Villages". *World Bank Research Observer*. **14**(11): 1-30.

Krishna, A. and Shariff, A. (2010). "The Irrelevance of National Strategies? Rural Poverty Dynamics in States and Regions of India, 1993-2005". *World Development*. **39**(4): 533-549.

Lanjouw, P. and Murgai, R. (2009). "Poverty decline, agricultural wages, and nonfarm employment in rural India: 1983-2004". *Agricultural Economics*. **40**(2): 243-263.

Ravallion, M. and Datt, G. (2002). " Why has economic growth been more pro-poor in some states of India than others?" *Journal of Development Economics*. **68**(2): 381-400.

CMI WORKING PAPERS
This series can be ordered from:

CMI (Chr. Michelsen Institute)
Phone: +47 47 93 80 00
Fax: +47 47 93 80 01
E-mail: cmi@cmi.no

P.O.Box 6033 Bedriftssenteret,
N-5892 Bergen, Norway
Visiting address:
Jekteviksbakken 31, Bergen

Web: www.cmi.no

Price: NOK 50
Printed version: ISSN 0804-3639
Electronic version: ISSN 1890-5048
Printed version: ISBN 978-82-8062-458-1
Electronic version: ISBN 978-82-8062-459-8

This paper is also available at:
www.cmi.no/publications

INDEXING TERMS
Poverty trap
Occupational choice
India

We investigate whether historic land distribution determines stagnation or development of Indian villages. The empirical analysis is motivated by the Banerjee and Newman (1993) model of occupational choice and economic development. Family histories are collected for a random sample of 800 households. Households are classified into economic categories according to the assets-occupations mix at present and at grandfather's time. Transitions are described, and for a remote district explained, by the historic village land distribution. We also investigate the role of social identity, and find that scheduled tribes are more likely trapped in poverty than scheduled castes.

CMI CHR. MICHELSEN INSTITUTE