

INTERGENERATIONAL ECONOMIC MOBILITY IN RURAL BANGLADESH

M Niaz Asadullah*
SKOPE
University of Oxford
Department of Economics
Manor Road
Oxford, OX1 3UQ, UK.

* I am grateful to Professor Stefan Dercon for his guidance. I would also like to thank seminar participants at the Centre for the Study of African Economies (Oxford) and the University of Reading for their helpful comments and suggestions. E-mail: mohammad.asadullah@economics.ox.ac.uk

Abstract

Unique family history data with retrospective information on parental assets are used to study socio-economic mobility in 141 villages in Bangladesh. We use transition matrix and regression based approaches to study mobility in household wealth across generations of the same family. An aggregate index of household assets is used to measure wealth. Regression estimates of father-son correlations and an analysis of intergenerational transition matrices show substantial persistence in wealth. We do not find wealth mobility to be higher between periods of a person's life than between generations. Lastly, various reasons for the lack of intergenerational mobility in wealth are explored. We find that the process of household division plays an important role: sons who splinter off from the father's household experience greater (albeit downward) mobility in wealth. Despite significant occupational mobility across generations, its contribution to wealth mobility, net of human capital attainment of individuals, appears insignificant. Low wealth mobility in our data is primarily explained by intergenerational persistence in educational attainment. A comparison of father-son schooling correlations for different cohorts further reveals that schooling mobility has not increased much in rural Bangladesh. We conclude by discussing the policy implications of these findings.

Key words: Household partition, intergenerational poverty trap, mobility indices, schooling mobility, transition matrices.

JEL classifications: I21, J62, O12.

Intergenerational Economic Mobility in Rural Bangladesh

1 Introduction

Prevalence of income poverty is a hallmark of countries at the bottom of the development discourse. As part of the concerted effort to reduce poverty, large scale investment in public infrastructure is undertaken in these countries, with an aim to equalize socio-economic opportunities. These investments are usually informed by periodic studies on poverty in developing countries that use cross-section data. Such studies nevertheless say little about the evolution of poverty and its long-term drivers and at best yield a static picture of the distribution of income. Consequently, they add little to our current understanding of the efficiency of various public investment programs. A significant part of the observed income inequality could reflect persistent differences in the capacity of individuals to exploit market opportunities. Economic immobility – persistence in economic status - then reflects inequality of opportunities and therefore the process that underlies persistence in (income) poverty. As such, two countries with an identical distribution of income in a given year can offer a very different set of economic opportunities to their populations and consequently differ in their capability to minimize income inequality over time and across generations. Therefore, in addition to growth in income and its distribution, knowledge of the extent of relative social mobility is useful in that it provides information about the long-run distribution of outcomes and the factors underlying them.

Despite the policy relevance of research on economic mobility, relatively less is known on the issue for developing countries. A handful of studies nevertheless exist on intra-generational (i.e. inter-temporal) mobility in rural economies of Asia using household level panel data. For example, Swaminathan (1991) uses such data from a South Indian village for the years 1977 and 1985 to examine mobility in wealth. Similarly, Fuwa (1999) uses data from the Philippines to study occupational mobility. In a much-publicised study of a north Indian village, Drèze *et al.* (1998) examine income mobility over five decades. However, study on persistence in outcomes across generations of the same family is rare.

A key reason for the absence of research on intergenerational mobility in developing countries is a lack of panel data or cross-section data with life histories of parents. Recently a number of studies have exploited the latter to study persistence in economic outcomes across generations. For example, Emran *et al.* (2003) examine occupational mobility using Nepalese data. They find that mother's participation in non-farm work substantially raises the probability of daughter's non-farm participation in Nepal. A more recent and comprehensive study on economic mobility is by Grawe (2004) who studies father-son earnings data from the US, the UK, Pakistan, Peru, Nepal, Malaysia and Ecuador. Grawe reports substantial earnings immobility in developing countries¹. Most importantly, when compared to developed countries, mobility is less in developing societies.

¹ Using a methodology and data similar to that of Grawe, Dunn (2003) reports low earnings mobility for Brazil. Fields (2000) discusses additional studies that use panel data from Peru and Malaysia to assess mobility in earnings.

This study investigates mobility - changes in the relative economic and social position of individuals - in rural Bangladesh over three decades using cross-sectional data with retrospective records on parental/household characteristics and asset portfolio. For 141 villages in Matlab thana, we examine the evolution of wealth for a sample of male household-heads across time and generations. While we do not study earnings or income mobility, through simultaneous analysis of mobility in wealth, education, and occupation we provide a detailed description of the dynamics of socio-economic mobility in a developing country. Our analysis suggests limited intergenerational wealth mobility in rural Bangladesh. Intergenerational persistence is also very high in educational attainment. Examination of various potential economic and demographic correlates of son's wealth shows that education is the most important driver of wealth mobility in rural Bangladesh.

The balance of the paper is as follows. Section 2 describes the study area – the Matlab villages - with reference to various aspects of the structural change that has transformed the Matlab area in the last three decades. This helps in explaining the societal context in which we study mobility. The empirical strategy is set out in section 4. It is then followed by a description of the data in section 5. Results are discussed in section 6 while section 7 concludes.

2 Social and Economic Backgrounds of Matlab, 1974-1996

The Matlab thana is located in Chandpur district, 55 kilometres south-east of Dhaka, the capital of Bangladesh. There are 141 villages in the area. The economy of Matlab is primarily agricultural, with the majority of population engaged in various farm activities. Despite its proximity to the capital, the villages remain considerably remote chiefly due to poor transport facilities. Travel between Matlab thana centre and the villages is primarily by foot/rickshaw/boat (Razzaque and Streatfield, 2001). The area is flat and low-lying. Being criss-crossed by rivers and canals, it is frequently flooded particularly during the monsoon. Rivers provide the most convenient route of transportation. Thus, despite its proximity to the capital, average travel time between the Matlab thana and Dhaka is five hours. Villagers therefore significantly rely on the local economy for their daily livelihood².

The period of 1974-1996 has seen notable socio-economic changes in Bangladesh which is also reflected in the data on the Matlab villages. By virtue of a comprehensive demographic surveillance system (DSS) maintained in the area since 1966 by the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), it is possible to track some of these changes. The extent of poverty and the exact changes in the standard of living over the years in the study area is difficult to ascertain in the absence of income/consumption data. Information on basic assets owned by the households, quality of housing, access to safe drinking water, and sanitation nevertheless exist in the DSS records. Three complete censuses conducted by ICDDR,B yield these data for the years 1974, 1982 and 1996. Analysis of this data suggests an improvement in material well-being and hence, perhaps a reduction in disparity in living conditions of the poor and non-poor households as well.

Substantial growth is observed in the ownership of productive assets and consumer durables (such as cows, boats, radios, televisions, quilts, lamps and so on) in the region.

² However, a fraction of the adult household members often work abroad as migrant workers and send remittance on a regular basis.

For example, household's ownership of radio has jumped from 10% to 40% while ownership of hurricane lamps has increased to 90% from 60% between 1974 and 1996. Similarly, there has been considerable improvement in materials used for construction of dwellings. The use of tin as a material for construction of roof increased considerably from 78% to 96%. In 1996, 38% of households used tin (or tin mix) to construct walls against 28% in 1974.

Additional gains in well-being are likely to have originated from an improvement in the public provision of safe drinking water. The vast majority of the population in the riverine villages of Matlab thana used to drink water from ponds, canals and rivers which was the principal cause of water-borne diseases such as Cholera and Diarrhoea in the 70s. In 1974, only 25% had access to drinking water from tubewells. However, provision of tubewells became almost universal over time so that 95% of the population secured access to "safe" drinking water from such wells by 1996³. This is thought to have significantly reduced mortality rate in the region.

It should be pointed out that detailed data on the evolution of landholding - the most important source of wealth in the rural agrarian societies - is unavailable. A snapshot of the distribution of land, however, can be obtained from data for the year 1996 (Appendix Table 1a). As many as 51% of the households are functionally landless (defined as owning arable land of less than .20 acre). Some increase in landlessness has occurred due to loss of land caused by river erosion⁴.

The precise mechanism underlying the changes in living conditions, as measured in terms of basic (non-land) indicators of household wealth, are not clear. "Gains" in living standard are apparently mirrored by a drop in illiteracy rate, a decline in fertility and changing occupational class structure. The rise in human capital level and the reduction in dependency ratio together with reduced dependence on farming (and increased employment in the service sector) could all potentially explain "improvements" in household well-being. We briefly discuss these three changes next.

Education: Table 1 provides the breakdown of the Matlab population in terms of their educational attainment. In the early 1970s, educational profile of the Matlab villages was not that different from rest of the country. The vast majority of population in Matlab were uneducated, with a staggering 72.7% of the females never going to school. A secular increase in school participation is nevertheless apparent since then. While the situation improved significantly for the males by 1982, low schooling remained a defining characteristic of the female population in the Matlab villages. Not only participation rate remained low, attainment was also limited: only 3.5% females had completed grade 7 or over by 1982 compared to 14.4% of males.

³ For washing and bathing purposes, however, most of the households continue to rely on ponds and river water.

⁴ These landless households mostly emanate from 7 villages that existed in 1982 but were subsequently lost following river erosions. The householders consequently relocated to neighbouring villages. But in general, studies for other regions of Bangladesh indicate a moderate decline in landownership over time. For example, Hossain et al. (2002) report 47% functionally landless households in 1987-88 against 49% in 1999-2000.

Table 1: Distribution of population (7 years +) by schooling by census years

Education	All			Male			Female		
	1974	1982	1996	1974	1982	1996	1974	1982	1996
0	66.9	59.9	40.5	61.2	49	33.4	72.7	70.9	47.4
1-3	16	15.6	20.1	17.9	19	22.3	13.9	12.3	18.5
4-6	10.8	15.5	21	10.3	17.7	21.4	11.2	13.4	20.6
7-9	3.1	5.7	10.6	4.6	8.5	11.9	1.7	2.9	9.3
10+	3.3	3.3	7.5	6	5.9	11	0.5	0.6	4.2
N	127922	147098	175646	64656	73841	86218	63262	73257	89428

Source: Reconstructed from Razzaque *et al.* (1998). All figures are in percentage.

The structural factors underlying the history of low adult education in Matlab are unknown. One potential factor is the limited supply of schools across Matlab villages. Most of the individuals aged 40 years or over in 1974 attended schools in the pre-1938 era, a period which is marked by presence of very few schools. The region saw substantial growth in schooling facilities in the post-1938 years which, however, slowed down somewhat during 1947-1970 when majority of the current adult population were of school age. When disaggregated by primary and secondary school types, most of the growth in the total number of schools has occurred in the primary sector. Such stagnation in the supply of secondary schools is mirrored by low secondary grade completion among adults in the Matlab villages. In 1996, apart from the smaller ones, a majority of the villages in Matlab had at least one school. The distribution of these schools by their management is given in Appendix Table 1.

Fertility decline: The past two decades have seen a steep decline in the fertility rate in Bangladesh. While such decline has also taken place in Matlab, the timing and speed of the decline has been affected by the presence of a large scale family planning program in the region. Since 1977, half of the Matlab villages have been exposed to an intensive family planning program run by the ICDDR,B. In addition to a contraception program, the treatment area also benefited from the presence of a maternal child health services. On the other hand, individuals in the control villages had access to the conventional family planning services offered by the government. Total fertility rate declined from 6.9 in 1976 to 2.9 in 1995 in the treatment area whereas the control area saw a smaller decline (from 7.2 to 3.6) over the same time period.

Labour force composition: In recent years, noteworthy changes have also occurred in the sectoral composition of the labour force. Table 1a reports the distribution of primary occupation of household-heads (aged 8 years and over) in Matlab thana for various years. It appears that the share of waged work (agricultural, non-agricultural and service related) declined from 35.7% in 1974 to 26.6% by 1996. This is primarily driven by a reduction in the share of the agricultural labour from 18% of the workforce in 1974 to 5.1% by 1996. This is consistent with the fact that among self-employed individuals, the share of non-farm activities increased while that of farm-work declined⁵. The share of waged work in non-farm (NF) sector also increased from 17.7% to 21.5% in 1996. These changes in occupational structure are suggestive of an overall reduction in the share of agriculture in the village economy.

⁵ Share of lower productivity-end non-farm activities (such as rickshaw pulling, fishing, boatman and so on) however remained constant, 9.1% in 1974 to 9.6% in 1996.

Table 1a: Distribution of primary occupation of household heads in various censuses

Occupation	Census year		
	1974	1982	1996
<i>1. Self-employment</i>			
- Owner-worker	35.2	31.4	24.3
- Rent-sharecropper	1.9	0.5	3.4
- Business (established, small, others)	6.4	8.3	12.4
- Rickshaw puller	-	-	2.1
- Fishing (catch/sell)	4.9	5.0	5.7
- Boatman	3.1	2.2	1.3
- Cottage industry	1.1	0.4	0.5
<i>2. Wage-employment</i>			
- Agricultural (wage) labour	18.0	21.2	5.1
- Non-agricultural (wage) labour (Mill/skilled/unskilled)	13.0	11.0	15.2
- Skilled service (including social work and doctors)	4.7	6.4	6.3
<i>3. Miscellaneous</i>			
- Unemployed/beggar/disabled/retired/not working/student	4.5	2.0	5.0
- Housewife/housework	6.1	11.0	18.0
- Others	1.0	0.7	0.5
- Unknown	0.0	0.1	0.0

Source: Adapted from Razzaque *et al.* (1998) and Ruzicka *et al.* (1974).

The finding that relative size of NF sector is quite substantial and growing is consistent with the national picture for the same period⁶. However, this does not necessarily imply a reduction in poverty. Despite similar transformation of the rural labour force in Bangladesh, the level of per capita rural NF income did not increase (Mahmud, 1996). This puzzle is partly explained by the composition of the NF sector in Bangladesh. Most of the individuals may have moved into low-productivity self-employment type activities, which is fuelled by an increase in landlessness in the country⁷. However, there may exist significant heterogeneity within the NF sector. For example, using the Bangladesh household expenditure survey data, Sen (1996) argues that income growth in NF sector is higher than that in agricultural wage employment if NF wage sector is excluded.

The above features of structural change since the early 1970s have important implications for social mobility in the Matlab villages. To the extent that these changes translate into income growth, one may wish to know how that affected the relative position of individuals in the income distribution. Besides, the question of drivers of descends and escape remains. Given the substantial return to education in the waged sector in the Matlab villages, the shift in employment in favour of waged work (particularly in non-farm sector) and the recent increase in schooling among villagers altogether may have contributed positively to well-being of the population⁸. Indirect evidence for such premonition comes from a recent study by Sen (2003), a longitudinal study of 397 rural households in Bangladesh for the period of 1987-2000. Households escaping poverty are found to have greater human capital, superior initial land endowment and a richer portfolio of physical and financial assets.

⁶ Similar pattern is also documented in Hossain *et al.* (2002) who study, inter alia, occupational change in rural Bangladesh over 1987-2000.

⁷ For example, according to Labour Force Survey 1991, 50% of the landless individuals in Bangladesh were employed in NF sector.

⁸Return to an additional year of schooling in wage work in Matlab villages is 8% (Berman and Stepanyan, 2003).

Lastly, a caveat is in order because of our use of data for the year 1974 as a benchmark. Two adverse shocks are likely to have affected the Matlab economy in the first half of 1970s. In years immediately after the liberation war of 1971, most of the Bangladeshis allegedly experienced a drastic drop in their standard of living, mainly due to major disruptions in the economy and society caused by the war. The worst victims of this process were industrial workers, small peasants, agricultural labourers and low earning groups. Between 1970 and 1973, per capita real income declined and real wages of agricultural and industrial labourers went down. A further fall in living standard is likely to have occurred due to the famine of 1974. The 1974 famine was mostly a rural phenomenon which unfolded following the summer flood in June-July. It might have affected the population in two ways, first, through increasing rural-urban migration (in search of food) and second, through sale of assets and land to smooth consumption. However, 1974 census data on Matlab villages (used in this paper) were collected in the first quarter of 1974 and hence, neither affected by the famine nor flood.

4 Empirical Strategy

Most of the existing empirical studies on intergenerational economic mobility employ the model of regression to the mean to relate son's status to that of his father. The resultant regression coefficient on the father's status is taken as a measure of mean persistence in status across generations of the same family. Regression estimates are usually complemented by an analysis of transition matrices. The latter provides relative rankings and offers a way to assess mobility as movements instead of mean persistence. In this study, we adopt both these approaches to study wealth and schooling mobility. However, the analysis of occupational mobility remains restricted to the transition probability approach for reasons discussed later.

4.1 Regression Approach to Mobility

We want to estimate the following regression:

$$\text{Son's Wealth}_t = \beta_0 + \beta_1 \text{Father's Wealth}_{t-l} + \nu_t \quad (1)$$

where ν_t is a random error term and “ l ” stands for the lag length in the measure of father's wealth. The regression framework focuses on mean mobility where β_1 is the OLS estimate of the degree of generational persistence in wealth. However, equation (1) does not say anything about the drivers of persistence in wealth. To this end, we report additional regressions of son's wealth *incrementally* controlling for the following covariates: age and age-squared, education, family size, inheritance of (household) headship and occupation. This then yields an alternative estimate of persistence, α_2 , from equation (2). This strategy allows us to track changes in the coefficient on father's wealth following inclusion of each of the additional determinants of son's wealth.

$$\text{Son's Wealth}_t = \alpha_0 + \alpha_1 \text{Age} + \alpha_2 \text{Father's Wealth}_{t-1} + \alpha_3 \text{Education} + \alpha_4 \text{Headship} \\ \text{Inheritance}_t + \alpha_5 \text{Occupation}_t + \alpha_6 \text{Family Size}_t + \varepsilon_t \quad (2)$$

The OLS estimate of persistence is not robust to the possibility that father's wealth could be measured with error so that estimate of persistence is smaller⁹. Therefore, we also report IV estimates of wealth persistence where father's wealth is instrumented by the father's education and occupation¹⁰.

A problem also arises with the son's family size, which is often co-determined with his wealth and hence endogenous. Wealthier households, particularly those engaged in farm work, are likely to demand more children in particular if the labour market is imperfect. Thus, we additionally instrument *family size* by (i) the sex of first (eldest) child and (ii) whether the household is located in the treatment area¹¹. Individuals in the treatment villages had access to specialized family planning services offered by the ICDDR,B. Since villages were randomly chosen for this intervention, location in the treatment area is a priori exogenous to household characteristics. Availability of multiple instruments means that our IV models are always over-identified. We therefore carry out conventional tests for the validity of our instruments.

A key feature of our data is that we are able to determine whether a son's household is a split-off or directly inherited from his father. This allows us to investigate another potential cause of mobility in rural areas i.e. *household partition*. Once again, household division could be driven by resource scarcity or a lack of surplus in one's household of origin (i.e. father's household). If father's wealth is measured with errors, exogeneity of household inheritance status would be of suspect: inheritance status (for sons) could partly capture the effect of unaccounted paternal wealth. We test for this possibility by additionally instrumenting son's inheritance of headship. As excluded instruments, we use two retrospective measures of "within household inequality": birth order of the son and difference in the educational attainments among his siblings. Within household inequalities in human capital (and age) implies that there is a reduced benefit of joint residence and consequently, an increased probability of household partition (Foster and Rosenzweig, 2001; Rosenzweig, 2003).

4.2 Transition Matrix Approach to Mobility

Estimates of father-son correlations obtained from equations (1) and (2) do not measure mobility as a positional change in one's wealth distribution. As such, it conceals important patterns of mobility across the whole range of father-son status distribution. In certain circumstances, instead of mean persistence, one may be interested in the mobility experienced by people at the bottom of the wealth distribution. The transition matrix approach is superior in all these respects. A (Markov) transition matrix is a square matrix describing the probabilities of moving from one state to another in a dynamic

⁹ There are very few studies on wealth persistence that tests for such a bias. However, the evidence of a downward bias is evident from the literature on earnings mobility. A series of revisionist studies (e.g. Solon, 1992; Zimmerman, 1992) report a larger estimate of persistence in the range of 0.4 to 0.6 (compared to older estimates of 0.2), once measurement errors in earnings are accounted for (Solon, 2002).

¹⁰ Charles and Hurst (2003) instrument father's wealth by father's education. Similar 2SLS framework is used in Dunn (2003). In earnings regressions for sons, Dunn instruments father's earnings by father's occupation. Nonetheless, the exogeneity of such instruments is not a priori conspicuous. For example, father's wealth could additionally suffer from the problem of endogeneity if unobserved earnings endowment (such as innate ability) is contained in the error term in equation (2).

¹¹ See Angrist and Evans (1998) for an application of similar instruments based on family composition. However, due to early marriage of daughters and migration of sons, this variable is possibly measured with error for older households, leading to downward bias and consequently the problem of weak instruments. However, as we show later, this problem does not arise in our data.

system. In each row are the probabilities of moving from the state represented by that row, to the other states. Thus all the elements in a given row of a transition matrix add up to one. In an intergenerational context, the matrix yields the probability of sons reaching a particular status for a given status of their fathers.

The technique works by converting continuous status variables for fathers and sons into discrete ordered variables having same number of ordered categories. Members of each generations/periods are classified according to fixed categories such as equal-sized quantiles, with base-period quantile determining the row and destination-period quantile the column¹². The joint distribution is then parameterized by a $(n \times n)$ matrix M containing the transition probabilities. If economic status persists fully across generations/periods, M would be an identity matrix with all entries lying on the diagonal axis.

The presence of multiple sons as heads of independent households in the sample means that the same father appears several times in the data when transition matrix or a regression model is estimated. This distorts the quintile position of fathers. Hence, we construct the fathers' distribution by counting each father only once and calculating his quantile ranking. The same procedure is followed when residuals from regressions are used to adjust raw wealth data for various covariates (such as age).

A key problem with the transition matrix analysis is that it is compounded by life-cycle effects. To this end, we use age-adjusted data instead, using residuals from a regression of wealth on individual's age and age-squared. In addition, we report additional transition matrices using residuals from regressions of son's wealth on the following sets of covariates: (i) age, age-squared, and education (ii) age, age-squared, education and family size and, (iii) age, age-squared, education, family size and occupation respectively¹³. A comparative analysis of these transition matrices provides a crude way to understand the factors that cause father-son mobility in the data. Therefore, in addition to the full sample, we estimate the transition matrices and associated regressions for a sub-sample of households that are inherited by sons and which are split-offs.

Given that the transition matrix approach is purely descriptive, it's difficult to compare mobility between two samples without some overall summary statistics or scalar measures. Therefore, for each transition matrix, M , we report two statistics which give a measure of mobility in terms of time dependence. These are Pearson's chi-square and likelihood-ratio chi-square statistics. These statistics compare expected frequencies (when there is perfect mobility) with observed frequencies and therefore assume larger values the further we are from the state of perfect mobility. Thus a matrix with a larger chi-squared value is said to be more time dependent and less mobile in the sense of *time dependence* than the other.

The above two indices, while useful for ranking transition matrices, do not say anything about *positional movements*: An individual is said to have experienced mobility if she changes position in the status distribution. Hence, for each transition matrix, we also construct an additional total of five indices of mobility: Shorrocks' MET (also known as the Prais index), Atkinson *et al.* mobility ratio, determinant index, average jump and normalised average jump.

¹² In case of occupations, such classification arises naturally.

¹³ To guard against possible outliers, we also repeat the analysis using residuals from median regressions.

These five mobility indices can be classified as (a) individual cell-related and (b) aggregate measures (Swaminathan, 1991). Both average absolute jump and normalised average jump are of the first type. The rest - Prais index, Atkinson *et al.* mobility ratio and the determinant index - are aggregate measures of mobility. The Prais index is defined as:

$$[\mathbf{n} - \text{trace of } \mathbf{M}]/[\mathbf{n}-1]$$

where, \mathbf{M} is the transition matrix and \mathbf{n} is the number of rows/columns. Atkinson (im)mobility ratio focuses on the fraction of cases lying along the principal diagonal and the adjacent cells. For a 3x3 transition matrix, for example, it is equivalent to $1/3*[(m_{11}+m_{21})+(m_{21}+m_{22}+m_{32})+(m_{23}+m_{33})]$, where m_{ij} is the ij -th element of the matrix. By focusing on the diagonal elements, these two measures provide a way to quantify the extent of immobility¹⁴. They vary with (i) number of quantiles and (ii) distance between initial and base year. The longer the time period, the smaller is the immobility ratio. The determinant index is defined as:

$$1 - |\mathbf{M}|^{1/[\mathbf{n}-1]}$$

Both the Prais index and the determinant index are bounded between 0 and 1: if there is perfect immobility, these two indices converge to zero.

The above three aggregate indices give no indication of how many quantiles a mover moves. In this respect, the cell-related indices -- mean absolute and normalized jump -- are superior: they provide a measure of the number of quantiles that the typical member of a class would jump between two periods. For example, “average absolute jump” calculates the mean number of quantiles moved in absolute value. Like the previous indicators of mobility, these two indices are also sensitive to the choice of the range over which movement is measured i.e. quintiles, ventiles and so on (Fields, 2000).

It should be noted that as an alternative to the positional change approach discussed above, we can examine relative mobility in terms of an individual’s share of total income/wealth vis-à-vis others. Here, instead of the position in the income distribution, the focus is on variations of income shares between the base and the final periods (Fields, 2000). This approach is sensitive to any small income changes within a quantile and the resultant change in the relative position one may experience. We therefore additionally report the correlation coefficient, a measure of share movement¹⁵. Fields further discusses two additional aspects of mobility: non-directional income movement and directional income movement. Based on an axiomatic approach, a measure of total mobility is derived which is decomposable into mobility arising from (i) transfer of wealth among individuals keeping total wealth constant and (ii) a change in total wealth available. As such, total mobility is decomposed into relative and resource mobility. However, these are not pursued in this paper. Consequently, we do not elaborate these conceptions any further.

¹⁴ However, measures of mobility with a focus on proximity to diagonal axis it not wholly unproblematic. For example, significance of a jump of one quintile depends on the location in the distribution - end quintiles are limited in their movements (Atkinson *et al.*, 1992).

¹⁵ It is also a measure of time dependence. The correlation coefficient is easily interpretable and intuitive but sensitive to outliers. In contrast, a rank correlation coefficient is robust to outliers but merely a measure of positional movement.

Lastly, an important caveat is in order. At present, there is no single measure of relative mobility that unambiguously ranks transition matrices: the ranking is often sensitive to the choice of mobility indicators (Valentino, 1993). Besides, most of the existing indices do not necessarily imply greater mobility for the transition matrices that reward individuals with poorer prospect in the base year/group a better prospect. While we estimate all six indices, for the sake of brevity, only four are used throughout the paper. These are: correlation coefficient, Prais index, Atkinson *et al.* mobility ratio and average jump index. This omission (of two indices) has no implication for our analysis as all the six mobility indicators were found to yield a nearly consistent ranking of the transition matrices for our data.

5 Data and Construction of the Working sample

The data on 141 villages of the Matlab thana used in this study comes from the ICDDR,B which has maintained a demographic surveillance system (DSS) in the area since 1966. We use a random sample extracted from the database on the Matlab Socioeconomic Census (MSEC) 1996, a complete census of the study villages carried out by the ICDDR,B. The sample is drawn using the following rule. First, we selected a random sample of 2687 bari from a total of 7440 bari in the Matlab thana. These 2687 bari are the same as those sampled for the Matlab Health and Socio-economic Survey (MHSS) 1996, an independent cross-section survey on the Matlab villages¹⁶. Then we extracted information on individuals residing in a total of 14868 households in the sample bari.

The MSEC 1996 sample data does not automatically yield past records on parental characteristics (such as age and education) and outcomes (such as occupations and asset portfolio). Rather, it is available retrospectively from earlier rounds of the MSEC if parents co-resided with their adult children in the same household in the past. Using information on relationship to household-head in earlier census records, one can reconstruct parental work history and other relevant characteristics. To this end, we bring in retrospective (socio-economic) information on parents and complete life history data of current household-heads in the following manner. We extract retrospective records on all individuals who shared a household in 1982 and/or 1974 with our 1996 sample individuals¹⁷. For a total of 12651 heads of households in our MSEC 96 sample, we are able to extract retrospective information implying an attrition rate of 14.9%. For the sample of male-headed households, the attrition rate is somewhat smaller i.e. 13.2% (see Appendix Table 1b).

It should be noted, however, that the randomness of our working sample due to such attrition is not seriously comprised. The characteristics of the included and excluded households (due to missing retrospective data) are strikingly similar. The few noticeable differences are: the heads of the excluded households are more educated, more often located in a single-household bari and have less cultivable land¹⁸. Appendix Table 1d summarizes these characteristics. However, an additional source of attrition prevails in the presence of non-random household division: more able/educated sons split and migrate outside the study area. Consequently, household residence (and/or residential cluster) based sampling akin to our sample may not yield unbiased estimates of economic mobility (Rosenzweig, 2003). We can only study mobility on the basis of life-histories of siblings who have continued to reside in the sample area. This remains a major limitation of our data.

For our empirical analysis, we construct two analytical samples, both consisting of male household heads¹⁹. The *first sample* consists of all heads (among the 10430 male heads in the MSEC 1996 sample) whose fathers were also present in the study area as a head in 1974. As such, we could obtain a random sample of father-son pairs with complete

¹⁶ The MHSS 96 chose bari as the primary sampling unit (PSU) rather than households and hence provide a better representation of family networks. Conditioning our sampling choice on the MHSS 96 sample bari also allows us to link our data to the latter.

¹⁷ The ICDDR,B collected census records for the entire Matlab population for the years 1974 and 1982.

¹⁸ The finding of residence in one-household bari is reassuring in the following sense. These (excluded) households are most likely to be recent migrants in the study area so that no information on their household of origin is available in the earlier MSEC records.

¹⁹ The focus on heads is because, data on assets is available only at the household level.

information on them and their household characteristics permitting a study of intergenerational mobility. A total of 5113 sons are identified in the MSEC 96 for whom their fathers were present as household heads in 1974. The remaining 5317 heads (in 1996) were discarded for any of the following four reasons: (1) the individual was also a head in 1974 (N= 4048); (2) the individual was not present in the household in 1974 so that relation to head could not be ascertained (N= 263); (3) the individual lived in the study area in 1974 but his household was headed by his mother instead (N=165); and (4) the individual was present in a male-headed household in 1974 but not related to the head as a son (N=775).

These conditions were imposed because records on parents were obtained by, first, locating parents as heads in the earlier census rounds and subsequently, linking individuals/heads as father and son through relation of the individuals to the head of their households in 1996 and 1974. It should be noted that further attrition in the data due to application of these rules does not lead to a loss of randomness of our sample apart from the second rule. However, rule 2 leads to a negligible reduction in sample size (i.e. a total of 263 observations) and hence not a serious concern. There is a further but small loss of observations (N=69) due to missing data on wealth for some parents/sons altogether resulting in a sample of 5044 sons for whom we have complete contemporaneous data on their households and retrospective information on their parents and household of origin in childhood.

5.1 Identifying the Split-offs

The information on whether a household headed by a son is a split-off is not recorded in the data. It is nevertheless possible to decompose our main sample of household heads (5044 sons) on the basis of the history of their households' formation. To this end, we follow Foster (1993) and apply 3 rules based on changes in the relationship of individuals to the household-head over time:

- i. If the *father is currently present* as the head of another household in the study area in 1996, the son's household is a split-off.
- ii. If the *father is present* as a non-head in a household in 1996, that household is inherited by the son who heads the household. Any remaining son observed to head a household where the father is not residing represents a split-off.
- iii. If two (or more) siblings are present as heads in 1996 and the *father is absent*, the eldest brother has the inherited household and the younger brother heads a split-off.

Application of these rules leads to a total of 2425 (47%) split-offs against 2688 (53%) inherited households²⁰. Given that we treat bari as the PSU in this study, our sample provides a better representation of the split-offs. This is because, anecdotal evidence suggests that sons who do not migrate outside their village of origin in most cases set up households in close proximity to father's household and hence located in the same bari.

5.2 Construction of Wealth Index

A major limitation of the Matlab censuses is that they neither contain data on value of the household assets nor is any information available on the stocks of assets reportedly

²⁰ This is comparable with Rosenzweig (2004) who, using Bangladeshi panel data, find that the rate of household division is 48% over a period of 18 years.

owned by a household. Information on sources of drinking water, quality of housing and various consumer durables nevertheless exist. In the absence of a unique measure of (long-term) wealth - expenditure data or value of household assets - we construct an aggregate measure of household wealth by combining data on household assets, quality of dwelling, usage and sources of water. These variables are: roof of largest room made of tin, wall of largest room made of tin, ownership of cows, boat, radio, clock, quilt and usage of water (i.e. drinking, cooking, and bath) from tubewell. The main challenge in creating a single index of household wealth (by aggregating qualitative/indicator variables) is the choice of appropriate weights. A popular solution to this problem is obtained following the principal component analysis (PCA). Almost all the asset variables are qualitative and naturally dichotomous so that PCA is a convenient procedure to obtain quantitative representation of household wealth based on individual indicators.

We implement the PCA only as an alternative procedure of aggregation²¹. Our preferred method of aggregation is one where weights for each item entering the household wealth index are derived from an underlying regression that explicitly links total value of household wealth to various assets. To be precise, using data on the linked MSEC-MHSS sample households, we regress total (log) value of household assets on the 12 variables additionally controlling for household size, schooling of head, and his spouse²². Data on asset value are obtained from the MHSS 96. The regression coefficients are reported in Appendix Table 2. The OLS coefficients on the 12 variables are stored and applied respectively as weights to aggregate the indicator variables into a scalar quantity, subsequently for various rounds of the MSEC data. An alternative set of coefficients is also derived by regressing household per capita expenditure on the same 12 variables with additional control for schooling of head and his spouse. The resultant coefficients yield an alternative set of weights and subsequently measure of household wealth for our data. However, like PCA based wealth index, we use this merely as a robustness check.

The use of regression coefficients as weights implies that the parental wealth variable in our study is a *generated one*. Use of such variable further in regression analysis does not bias estimated regression coefficients. But statistical inference is likely to be compromised. The exact direction of the bias in the standard errors is not clear a priori. If the bias is downward, t-statistics are likely to be inflated and hence problematic. We examine this issue later in the empirical section by bootstrapping the standard errors.

The sample characteristics along with descriptive statistics of the wealth indices are reported in Appendix Table 1c. Summary statistics are also reported separately for the sample of split-offs and inherited households. Compared to sons who inherited the headship, split-offs have less wealth, headed by less educated and younger males, and have fewer individuals in the households. Heads of split-offs are more into non-farm self-employment type activities whereas inherited heads are found more in self-employment in the farming sector.

²¹ See Filmer and Pritchett (2001) for an application of this technique. There is no consensus on the choice of the number of factors. We followed the existing practice and chose the first factor.

²² Total value of household assets comprised of current value of the followings items: homestead land, ornaments, savings, television, radio, clock, fan, bicycle, furniture and quilt.

6 Results and Discussion

6.1 Intergenerational Wealth Mobility

6.1.1 Regression Estimates of Persistence in Wealth

Appendix Table 3 reports regression estimates of (mean) persistence in wealth for the full sample. The OLS estimate of father-son persistence falls from 0.53 (column 1) to 0.35 (column 2) as we additionally control for son's age, education, family size and occupation. However, the largest fall occurs when we add the education variable. Despite significant occupational transitions observed in the Matlab villages over the last three decades, occupational mobility does not seem to have any impact on wealth mobility. To test this more explicitly, we re-ran column 2 specification (with and without control for son's education) additionally controlling for father's occupation so that the coefficients on son's occupations would capture the effect of occupational change across generations (results not shown). In neither of the two experiments, the coefficient on father's wealth changed significantly confirming that occupational mobility did not contribute to wealth mobility in our data.

It may be recalled that inclusion of father's wealth and son's family size are problematic due to the problem of measurement error and the endogeneity respectively²³. To examine the bias in our regression estimate of wealth persistence, model 3 jointly instruments father's wealth and son's family size. Excluded instruments are father's education and two occupation dummies (indicating participation in self-employment in agriculture and non-agriculture). The instruments for family size are (i) a treatment area dummy and (ii) a dummy indicating whether sex of first child is a son²⁴. Instrumenting father's wealth always leads to a significant increase in our estimate of persistence in wealth. Durbin-Wu-Hausman test statistics comfortably rejects the exogeneity of father's wealth in all cases. This finding tends to support the view that OLS estimates are possibly downward biased due to the presence of measurement error.

However, given the nature of our excluded instruments, it is not possible to be conclusive about the exact source of any potential bias. The use of occupation and education as excluded instruments for father's wealth implies that the IV estimate of the coefficient on father's wealth reflects father-son persistence in the *earned income* or measured portion of wealth (as predicted by his human capital and occupational choices). Like father's wealth, these two instruments remain potentially correlated with unobserved earnings endowment (such as innate ability) that is common between fathers and sons. If true, father's wealth is endogenous and the reported IV estimates are unlikely to be robust to such problem.

Interestingly, the inclusion of the son's family size variable has no effect on the size of the coefficient on father's wealth, even when we treat family size as endogenous. The latter has a significant effect on son's wealth (+ve in OLS specification and -ve in IV). Durbin-Wu-Hausman test statistics comfortably rejects the exogeneity of family size in all cases at 5% level. The excluded instruments are always highly significant in the first

²³ Although, as pointed out before, father's wealth could be additionally endogenous.

²⁴ We estimated separate models where we individually treated father's wealth and son's family size as endogenous. In both cases, our instruments comfortably passed the over-identification test. However, for the sake of brevity, we have suppressed these results.

stage regressions and pass the validity test in 8 out of 12 cases²⁵. Model 5 and 6 further expand the regression specification with control for inheritance of household headship. Model 7 additionally instruments the headship inheritance dummy by two (retrospective) measures of within household inequality.

Models 8-13 in the Appendix Table 3 extend models 2-7 with additional control for bari fixed effects. These fixed effects wipe out correlation between households in the bari that were joint in recent past and therefore, may have shared significant economic ties. As a further robustness check, we repeated our analysis with additional indices of household wealth, constructed following the PCA and weights derived from expenditure regression (results suppressed). However, our findings go through. It may be recalled that father's wealth variable is a generated regressor in our model so that inference may be incorrect. This is particularly a problem if t-statistics are marginally significant and standard errors are biased downwards. The t-statistics reported in Appendix Tables 3 and 4 are very large for all of the key variables of interest, particularly father's wealth. We nevertheless test whether our inference is compromised by bootstrapping the standard errors (results not shown). However, our earlier conclusions hold; while bootstrapped standard errors in some cases yield smaller t-statistics, they still remain significant at the conventional level.

Similar results are reported for sons who head split-offs and inherited households (Appendix Table 4). Simultaneously correcting for the measurement error in father's wealth and endogeneity of son's family size and controlling for all other covariates, the resulting estimates suggest higher persistence for sons heading inherited households. However, it is not known whether the difference in persistence with the split-offs is significant at conventional levels.

6.1.2 Transition Matrices of Mobility in Wealth

Appendix Table 5 reports the wealth (quintile) transition matrices for the sample of father-son pairs. Sons are represented in the columns against fathers' rows. For each son, the matrices give us the probability of being in a certain wealth quintile conditional on father's position in his wealth distribution. Hence, the row probabilities add up to 1. On the basis of the raw data, movement appears to be restricted to individuals who are initially better-off (in terms of parental position). The chi-square test statistics have large values (all significant at 1% level) indicating substantial time-dependence. However, without correction for age of the individuals, these matrices are confounded by life-cycle effects²⁶. Indeed the age-adjusted transition matrix exhibits greater mobility compared to that for the raw data. Substantial persistence remains nevertheless, particularly at the two ends of the wealth distribution. The probability that sons of the poorest stay poor is 0.34. Likewise, sons of the richest fathers stay richest in 36% cases. In general, immobility is much higher in the highest and the lowest wealth quintiles than in the middle.

In Appendix Table 5, we additionally compute three transition matrices using residuals from the OLS regressions with control for key determinants of son's wealth. To be precise, we incrementally adjust son's wealth data for his educational attainment, occupation and household size. Comparison of the age-adjusted transition matrix to

²⁵ However, they pass in all cases only if IV regressions with endogenous family size and un-instrumented parental wealth are considered.

²⁶ Given that data on fathers and sons are from 1974 and 1996 respectively, this problem is somewhat less serious in our data.

these latter matrices provides a crude way to understand how persistence in education and occupation and shocks to fertility over time may have affected mobility in wealth. For example, convergence of (transition) cell probabilities to the neighbourhood of 0.2 with an additional control for, say, education, would suggest that lower educational mobility is associated with immobility in raw/age-adjusted wealth data.

Indeed a large number of the cell probabilities converge to values in the neighbourhood of 0.20, particularly on and around the principal diagonal. The largest drop in cell probabilities, particularly at the extreme end of the wealth distribution, occurs when we adjust for son's educational attainment. This is also evident from the drop in chi-square values of likelihood statistics, a measure of time dependence. The relatively marginal drop in cell probabilities following additional control for son's occupation is not surprising given that we already control for education and it is the human capital development of sons that may have facilitated occupational mobility from farm to non-farm activities, particularly for the poor fathers²⁷. Hence, of all the characteristics, son's education appears to be the key to mobility in wealth.

However, as pointed out earlier, it is difficult to compare the degree of relative mobility for different samples only on the basis of the underlying transition matrices or indicators of time dependence (such as chi-square tests). Some corresponding summary statistics (of positional movement) are required for the purpose of ranking the matrices. We therefore turn to indices of mobility presented in Table 4. Column 1 reports the indices that correspond to transition matrix for raw wealth data. Columns 2-5 incrementally adjust son's wealth data for various covariates of wealth. Altogether, these indices correspond to the transition matrices reported in Appendix Table 5. As we move across the columns from left to right, there is an increase in the value of the indices for all samples. The observed increase in mobility thus highlights the underlying forces for initial persistence in wealth in raw data.

Table 4: Indices of wealth mobility, father-son pairs

		1	2	3	4	5
Full sample	Correlation coefficient	0.507	0.432	.300	0.303	0.286
	Prais index	0.847	0.882	0.933	0.931	0.939
	Atkinson <i>et al.</i> Mobility Ratio	0.262	0.307	0.354	0.354	0.361
	Average Jump	1.028	1.141	1.271	1.271	1.293
Inherited Households	Correlation coefficient	0.513	0.454	.310	0.312	0.297
	Prais index	0.831	0.867	0.934	0.931	0.935
	Atkinson <i>et al.</i> Mobility Ratio	0.253	0.298	0.352	0.354	0.363
	Average Jump	1.005	1.124	1.275	1.274	1.295
Spilt-off households	Correlation coefficient	0.501	0.408	.289	0.292	0.274
	Prais index	0.871	0.901	0.932	0.931	0.944
	Atkinson <i>et al.</i> Mobility Ratio	0.279	0.319	0.358	0.358	0.361
	Average Jump	1.071	1.164	1.272	1.272	1.294

Note: column 1 refers to raw data; column 2 uses age (and age squared) adjusted residuals from OLS regressions; column 3 uses age and education adjusted data; columns 4 & 5 additionally adjust for family-size and occupation respectively (using OLS residual).

²⁷ Hossain *et al.* (2002) make a similar observation in their panel study of 32 Bangladeshi villages. Households with primary earner in service sector had highest education for their workers, followed by those dependent on trade, business and farming, particularly agricultural wage work.

Table 4a: Indices of wealth mobility, father-son pairs

	1	2	3
Correlation coefficient	0.290	0.268	0.261
Shorrocks's MET - the Prais index	0.933	0.939	0.942
Atkinson et al. Mobility Ratio	0.357	0.373	0.369
Average Jump	1.282	1.309	1.311

Note: column 1 uses residuals from OLS regressions with control for age, age-squared, education, family-size, occupation and headship inheritance dummy. Columns 2 and 3 use residuals from IV regressions. Column 3 treats family size and headship inheritance as endogenous while column 2 only treats family size as endogenous.

To get a better idea about the extent of mobility displayed by various matrices, reported values of the indices need to be compared with values assumed under “perfect mobility”. The correlation coefficient ranges from 0 to 1 (no mobility) so that the value of .50 for the full sample implies that 50% of the variation in son’s wealth is attributable to his father’s wealth. When the transition matrix is defined in terms of deciles, the expected average jump in a state of perfect mobility is 3.3; where quintiles are used, this benchmark is scaled down to 1.65. Thus the range of “average jump” index from 1.02 to 1.29 for the full sample represents 64% to 96% of the value under perfect mobility. Similarly, when there is perfect mobility, all transition probabilities are equal to 0.2 so that Atkinson *et al* index, Shorrocks' MET and determinant index converge to a value of 0.52, 1 and 1 respectively. As evidenced in Table 4, Atkinson *et al.* index shoots from about 0.26 (age-adjusted data, column 2) to 0.36 (OLS residuals with age, education, occupation and family size adjustments, column 5). This is equivalent to 11% increase in mobility. Nevertheless, careful comparison confirms our earlier finding that control for son’s education leads to most of the gains in mobility.

We also implemented various robustness checks to verify the above results. To check the sensitivity of our results to measurement error, we also used residuals from median regressions²⁸. However, our results go through. We additionally controlled for inheritance of headship and tested whether the results were affected by the endogeneity of certain covariates (i.e. family size and inheritance of headship). To this end, residuals from IV regressions are used to compute the transition matrices for the full sample. Corresponding mobility indices are reported in Table 4a. Comparison of column 3 (Table 4a) to column 5 of Table 4 shows that apart from the correlation coefficient and the index of average jump, our results remain fairly robust.

Turning to results for various sub-samples, the same pattern holds for the sample of sons who head inherited households and split-offs. Initially, the mobility indices (computed using raw as well as age-adjusted wealth data) have slightly higher values for split-offs suggesting that moving out of parent’s household leads to greater mobility. However, once we additionally adjust wealth data for education, family size, and occupation, almost all the indices converge to same set of values (columns 3-5). This is particularly evident as we look at the transition probabilities in Appendix Table 5 for these three samples. After detailed control, a large number of the cell probabilities converge to a value in the neighbourhood of 0.2, the perfect mobility benchmark for cell probabilities. Once again, the largest fall follows from additional control for son’s educational attainment suggesting that, like the full sample, education remains the key drivers of mobility for the sub-samples.

²⁸ Compared to OLS, median regressions are more robust to outliers

6.1.3 Inter-temporal Mobility in Wealth

A total of 4048 male-heads in our 1996 sample headed the same household in the years 1974 and 1982, so that it is possible to study inter-temporal mobility for this group of individuals, particularly in wealth. For the sake of brevity, we do not report the transition matrices; only the corresponding indices of mobility are presented in Table 5. For this sample, the principal cause of mobility is life cyclical and emanates from the permanent income hypothesis. Individuals tend to low mobility earlier in life. Income increases with age, but at retirement it declines as savings are depleted and transfers are made to progeny. Like intergenerational mobility, we report transition matrix with corrections for age, as some individuals are likely to be in the middle of their life cycle in 1974.

Table 5: Indices of mobility in wealth, 1974-1996

	1	2	3	4	5
Correlation coefficient	0.539	0.526	0.389	0.385	0.378
Prais index	0.848	0.846	0.906	0.901	0.913
Atkinson et al. Mobility Ratio	0.267	0.262	0.33	0.331	0.346
Average Jump	1.046	1.041	1.192	1.191	1.229

Note: column 1 refers to raw data; column 2 uses age and age squared adjusted data (residuals from OLS regressions); column 3 additionally adjusts for education; columns 4 & 5 additionally correct wealth data for family-size and occupation (using residuals from OLS) respectively.

Table 5 shows substantial persistence in an inter-temporal context: individuals with poor initial wealth tend to be worse-off later in life. Given the gap of 22 years between the base period and the final period, time dependence is substantial. This is true even after adjusting for changes in the family size and controlling for an individual's age. Once we adjust for educational attainment, the probability of being on the diagonal axis is substantially reduced, perhaps implying that the lack of education also limits mobility during one's life cycle. This is similar to intergenerational mobility, where persistence weakens once such adjustments are made. Furthermore, a comparison of inter-temporal and intergenerational mobility indices (Table 4 vs. Table 5) suggests that persistence is not necessarily higher in an inter-temporal context.

In sum, intergenerational wealth mobility remains limited in rural Bangladesh. This is *prima facie* not so surprising given that we do not assess father-son mobility before transfer of bequests. However, the estimate of persistence is large considering the fact that we control for household-headship, the most prominent form of bequests in rural South Asia. A comparison of conditional and unconditional transition matrices suggest that education is the key determinant of wealth mobility in our data. Net of schooling attainment, however, occupational father-son differences has no impact on economic mobility. In the next two sections, we explore in detail persistence in occupation and education. In particular, we compare estimates of schooling persistence for two cohorts of adults and examine the evolution of schooling mobility in the study area.

6.2 Intergenerational Mobility in Occupation

Our analysis of occupational mobility is primarily based on occupational transition matrix analysis. We categorize individuals into five socio-economic classes on the basis of their “primary” occupations. As discussed earlier, individuals are primarily self-employed in various types of farm and non-farm activities in Matlab as owner-worker, rent-sharecropper, established and small businesses, rickshaw puller, fishing, boatman and cottage industry. Wage employees are either in agriculture or in non-farm sector as mill/skilled/unskilled workers or other skilled service-holders. As much as 30% of the sons in 1996 were engaged in wage employment with 23% being in non-farm wage work. Only a small fraction (1.3%) reports them as unemployed,

Table 6: Transition matrices of occupational choice, father-son pairs

		Son						
		1 (Self employment in agriculture)	2 (Self employment in non-agriculture)	3 (Wage employment in agriculture)	4 (Wage employment in non-agriculture)	5 (not working)		
Father	All	1	0.43	0.25	0.06	0.24	0.02	
		2	0.11	0.68	0.04	0.17	0.01	
		3	0.21	0.36	0.13	0.29	0.01	
		4	0.19	0.47	0.03	0.29	0.02	
		5	0.43	0.26	0.07	0.22	0.02	
	Inherited Households		1	2	3	4	5	
		1	0.46	0.22	0.06	0.24	0.02	
		2	0.12	0.68	0.02	0.16	0.01	
		3	0.26	0.36	0.11	0.26	0.01	
		4	0.24	0.41	0.03	0.3	0.02	
		5	0.44	0.26	0.06	0.21	0.03	
		Split-off Households		1	2	3	4	5
			1	0.4	0.28	0.07	0.24	0.01
			2	0.1	0.68	0.05	0.17	0
			3	0.17	0.37	0.14	0.31	0
4	0.14		0.54	0.03	0.28	0.01		
	5	0.41	0.26	0.09	0.24	0		

Table 6a: Indices of occupational mobility, father-son pairs

	All	Inherited	Split-offs
Correlation coefficient	--	--	--
Prais index	0.861	0.852	0.874
Atkinson <i>et al.</i> Mobility Ratio	0.427	0.434	0.421
Average Jump	1.434	1.449	1.417

Note: Correlation coefficient is not reported for discrete nature of the occupation data.

retired, disabled or not working²⁹. The extent of occupational persistence for all the categories becomes pertinent from Table 6 which reports the transition matrix. There is

²⁹ This group also includes few individuals who report begging as their occupation.

relatively less mobility among sons of self-employed individuals - in 43% and 68% cases, sons and fathers both were in farm and non-farm self-employment respectively. However, substantial movement out of agricultural labour is observed- in 29% and 36% cases, sons of agricultural wage labourers move into waged work and self-employment in non-farm sector respectively. This finding is very significant as this category is usually considered poorest of the poor. The wage-labourers usually have little land and small asset endowment. The majority appears to have moved to self-employment, particularly in non-farm work.

Our finding of higher persistence in the non-farm sector (self-employment and wage work combined) is in contrast to Emran *et al.* (2003). They find no effect of having a father in non-farm on son's probability of non-farm participation, particularly when they correct for the endogeneity of son's education and assets³⁰. A positive intergenerational effect is found only in skilled jobs. To the extent non-farm self-employment activities involve high skills in rural Bangladesh, this finding is consistent with our data. This point becomes clear if we disaggregate non-farm work of fathers into wage work and self-employment. Occupational persistence is greater among sons of individuals who were self-employed into non-farm activities. To this end, we experimented with a simple reduced-form Probit regressions of occupational choice where son's non-farm participation was regressed against education and two dummies for father's occupational status (self-employment and wage-employment in non-farm work). When son is in non-farm waged work, we find little (although significant) influence of father's occupation dummies (for non-farm engagements). On the contrary, replacing the dependent variable by a dummy for son's employment in non-farm self-employment work, we find large (and significant) coefficients on father's non-farm work status dummies. These results are consistent with the conclusion that follows from the occupational transition matrices: the large number of sons entering the non-farm sector reflects occupational mobility only when we focus on those who went into waged work.

6.3 Intergenerational Schooling Mobility

Given the large father-son mean persistence in schooling, it is of interest to study educational mobility as relative movements experienced by sons for a given level of schooling of their father. This can be accomplished using a transition matrix analysis. Since the distribution of schooling in 1974 was largely skewed (i.e. the majority of fathers being uneducated), we do not compute the quantile transition matrix. Instead, we classify individuals in five distinct groups on the basis of their levels of schooling. These are: no education, less than primary education (grade 1-4 completion), primary education (grade 5 completion), junior secondary education (grade 6-8) and secondary education (grade 9 and above). Table 7 reports the resultant transition matrices of schooling for the sample of father-son pairs. The corresponding indices of mobility/immobility are reported in Table 7a.

³⁰ However, their instruments for the schooling variable are rather weak in the first stage regression for the sample of sons. Hence it is possible that their results are reflecting a bias in the IV estimates due to the application of weak instruments.

Table 7: Transition matrices of schooling, father-son pairs

		Sons					
		No education	Less than primary	Primary	Junior secondary	Secondary or above	
		1	2	3	4	5	
Fathers	Full sample	1	0.59	0.18	0.09	0.06	0.08
		2	0.36	0.22	0.14	0.11	0.18
		3	0.28	0.21	0.16	0.15	0.2
		4	0.15	0.18	0.19	0.14	0.34
		5	0.11	0.12	0.07	0.15	0.55
		Pearson's χ^2					842.63
		Likelihood χ^2					791.00
	Inherited households	1	0.56	0.19	0.09	0.07	0.09
		2	0.32	0.23	0.15	0.11	0.19
		3	0.26	0.22	0.16	0.16	0.2
		4	0.13	0.17	0.19	0.17	0.35
		5	0.07	0.13	0.04	0.16	0.6
		Pearson's χ^2					514.98
		Likelihood χ^2					492.40
	Split-offs	1	0.61	0.17	0.09	0.06	0.07
2		0.41	0.22	0.11	0.1	0.16	
3		0.31	0.2	0.16	0.14	0.2	
4		0.18	0.21	0.19	0.1	0.31	
5		0.19	0.11	0.11	0.13	0.46	
	Pearson's χ^2					308.58	
	Likelihood χ^2					287.74	

Table 7a: Indices of schooling mobility, father-son pairs

	All sons	Inherited	Split-offs
Correlation coefficient	0.390	0.409	0.351
Prais index	0.834	0.822	0.862
Atkinson <i>et al.</i> Mobility Ratio	0.324	0.31	0.355
Average Jump	1.171	1.137	1.251

Intergenerational persistence is very high among sons of uneducated fathers- almost 60% of them stay uneducated whilst only 14% manage to obtain education beyond primary schooling. Interestingly, split-offs have less time independence than sons who head inherited households (as indicated by relatively smaller values of chi-square test statistics). This is consistent with the larger values of the mobility indices for split-offs indicating greater positional movement. Using the benchmark figures of these indices under the state of perfect mobility, it is possible to assess the relative mobility experienced by individuals in different samples. For example, the “average jump” index yields 68% mobility for inherited household heads compared to 75% for split-offs³¹. Similarly, The Atkinson *et al.* index yields 59% mobility for inherited household heads compared to 67% for split-offs³². The value of correlation coefficient is .39 for the full sample and

³¹ These figures are computed by comparing the raw values to the benchmark figure (under perfect mobility) of 1.65.

³² These figures are computed by comparing the raw values to the benchmark figure (under perfect mobility) of 0.52.

implies that 39% of the variation in son's education is attributable to variation in the father's schooling.

However for the split-offs, the direction of mobility is mostly downward. For example, the probability of descending from the top group to the bottom (i.e. staying uneducated when fathers have completed secondary education or above) is 0.19 compared to only 0.07 for the sample of inherited households. Our finding that sons who experience greater downward mobility in schooling are also heads of split-offs is consistent with Foster and Rosenzweig (2001). For Indian data, they report that sons who splinter from parents have lower school attainment compared to those who inherited (and hence headed intact) households.

Given the evidence of a lack of intergenerational mobility in schooling, one is interested to know whether schooling persistence has declined over time and its implications economic mobility. The relationship between schooling mobility and earnings (economic) mobility has been formalised in a simple model discussed in Solon (2004). This can be summarised using the following relationships:

$$\begin{aligned} \text{Earnings}_t &= \phi \text{Schooling}_t + \varepsilon_t \dots\dots[3] \\ \text{Schooling}_t &= \psi(\text{Parental income})_{t-1} + v_t \dots\dots[4] \\ \text{Earnings}_t &= \phi\psi(\text{Parental income})_{t-1} + \omega_t \dots\dots[5] \\ &[\omega_t = \phi v_t + \varepsilon_t] \end{aligned}$$

Equation (3) specifies earnings as a function of schooling attainment whilst equation (4) expresses schooling as a function of parental income. The parameters ϕ and ψ stand for labour market returns to education and elasticity of schooling with respect to parental wealth, respectively. If we combine equations (3) and (4) by expressing earnings as a function of parental wealth, equation (3) reduces to equation (5). According to Solon (2004), the intergenerational mobility parameter is now $\Phi\Psi$. This interpretation of the intergenerational correlation has the following implication. Earnings mobility is higher if schooling attainment is sensitive to parental and the returns to education are positive. Analysis of labour market earnings data suggests that, ϕ , the returns to education in Bangladesh is significant and ranges between 7% and 8% (Asadullah, 2006; Berman and Stepanyan, 2003). Therefore, one is interested to know the size and the evolution of Ψ . If sensitivity of schooling vis-à-vis parental income distribution turns out to be high and stable over time, this would undermine economic mobility in the future.

To answer these questions, Table 8 reports regression estimates of intergenerational persistence in school completion. The OLS regression of sons' schooling on that of their fathers (with no other covariates included) yields an estimate of 0.51 which drops somewhat with the inclusion of mother's education. The effect of maternal education always dominates that of paternal education. The coefficient on father's schooling reduces furthermore to 0.29 with the inclusion of father's wealth. The effect of parental wealth is also evident from the jump in adjusted R² values (from 0.17 to 0.27). However, no further changes occur to the coefficient on father's wealth as we additionally control for father's age, family size, occupation and village of residence. Hence, net of parental wealth and maternal education, the influence of father's schooling on that of the son remains large and significant.

Table 8: OLS estimates of intergenerational persistence in schooling [Dependent variable: son's school attainment]

	(1)	(2)	(3)	(4)	(5)	(6)
Father's education	0.515 (25.92)**	0.416 (18.50)**	0.290 (12.82)**	0.290 (12.72)**	0.285 (12.53)**	0.273 (12.75)**
Mother's education		0.512 (10.04)**	0.397 (8.20)**	0.395 (8.13)**	0.410 (8.44)**	0.369 (8.16)**
Father's wealth			1.897 (22.10)**	1.863 (21.55)**	1.767 (19.51)**	1.858 (20.04)**
Father's family size					0.011 (0.45)	0.006 (0.26)
Father's age					0.021 (4.16)**	0.022 (4.26)**
Father self-employed in non-agriculture				-0.439 (3.77)**	-0.390 (3.33)**	-0.300 (2.09)*
Father wage-employed in non-agriculture				0.013 (0.08)	0.098 (0.61)	0.250 (1.53)
Adjusted R ²	0.15	0.17	0.27	0.27	0.27	0.26
Village fixed effects	No	No	No	No	No	Yes
N	4150	4150	4150	4150	4150	4150

Note: Absolute value of t statistics in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%. All information apart from son's schooling corresponds to the year 1974. Sons are restricted to be of school age (aged 24 years or less) in 1974. Regressions include a dummy for missing data on mother's education. Fixed effect controls for village location in 1974.

From Table 8, it is apparent that Ψ is large and positive. Schooling attainment is very sensitive to parental wealth (and family background) in rural Bangladesh. It is therefore important to ascertain whether Ψ has declined over time. To this end, Table 8a reports estimates of schooling persistence using data on a sample of current household heads and their adult co-resident sons. Clearly, intergenerational persistence in schooling has remained largely stable when we consider correlation for a much younger cohort of school graduates. This is evidence if we compare the coefficients on father's wealth in Tables 8 and 8a.

Table 8a: OLS estimates of intergenerational persistence in schooling [Dependent variable: son's school attainment, aged 17-24 years]

	(1)	(2)	(3)	(4)	(5)	(6)
Father's education	0.442 (33.71)**	0.308 (18.89)**	0.216 (13.58)**	0.184 (12.25)**	0.184 (12.24)**	0.186 (12.36)**
Mother's education		0.331 (13.97)**	0.254 (11.23)**	0.211 (9.86)**	0.210 (9.79)**	0.219 (10.15)**
Father's wealth			1.946 (24.50)**	1.638 (21.49)**	1.646 (21.31)**	1.601 (20.52)**
Father's family size					-0.012 (0.64)	-0.015 (0.79)
Father's age						0.019 (3.81)**
Father self-employed in non-agriculture				-2.588 (22.73)**	-2.586 (22.71)**	-2.583 (22.71)**
Father wage-employed in non-agriculture				-1.697 (13.38)**	-1.696 (13.37)**	-1.703 (13.44)**
R-squared	0.18	0.22	0.30	0.38	0.38	0.38
Observations	4.019	3.964	0.604	2.124	2.194	1.267

Note: Absolute value of t statistics in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%. All information apart from corresponds to the year 1996. Sons are restricted to be of school age (aged 24 years or less) in 1974. Fixed effect controls for village location.

The fact that Ψ has remained stable over time suggests that much of the recent increase in school participation and attainment in rural Bangladesh has been unequally distributed across socio-economic groups. Returns to education are significant and positive and educational attainment helps escape poverty trap. However, given the inequality in educational expansion, sons who were born to uneducated poor fathers have not benefited from this expansion. This in turn has limited upward economic mobility and reinforced the process of intergenerational persistence.

7 Conclusion

This paper has looked at the nature and extent of socio-economic mobility among male household-heads in Matlab villages in Bangladesh. While we did not have access to a panel dataset containing repeated information on father-son pairs, we have used retrospective residential records on a sample of current household heads to track their household of origins. As such, we were able to fully reconstruct their family history with retrospective information on their parents. This has been possible by matching Census records on household-heads in the Matlab villages in 1996 to their parents on whom data was collected in an earlier census in 1974. By construction, the resultant dataset permits a study of mobility only among those sons who have remained in the study area over the last 22 years. Given that we chose bari as the primary sampling unit for the selection of households, our dataset contains most of the adult sons who currently heads individual households in the study area³³. However, the prospect of out-migration (particularly among “high ability” sons) implies that our data is censored. The sample attrition owing to non-random omission of split-offs (moving out of the study area) could bias the estimates of mobility and its various determinants (Rosenzweig, 2003). This possibility remains an important limitation of our study. With this caveat in mind, the following results emanate from this study.

We find large intergenerational persistence in raw (and age-adjusted) wealth data, particularly at the two tails of the wealth distribution. However, such persistence weakens once we additionally net out contribution of son’s education, household inheritance and family size to his wealth. The fall in transition probabilities (on the diagonal axis) is the largest when we adjust for educational attainment, indicating that differential schooling is the key source of persistence in wealth across generations of the same family. The cell probabilities remain largely stable once we occupational differences among fathers and sons are controlled for. Regression analysis of wealth data yields an estimate of father-son wealth elasticity in the range of 0.77 and 0.53. This is much larger than the existing estimates for developed countries and confirms the commonly held view that economic mobility is in developing societies³⁴.

An additional contribution of this paper is the construction of profiles of generational mobility for sons by inheritance of household headship. As such, we shed some light on a key demographic source of immobility in rural societies: household partition. While moving out of the household of the patriarch creates greater mobility in wealth accumulation for sons, we find that in most cases it is downward.

In addition to wealth mobility, we also assess mobility in school completion and occupational choice. Regression analysis of schooling data on father-son pairs exhibits significant intergenerational persistence. Educational transition matrices reveal that mobility is very low among children of uneducated fathers: almost 60% of them stay uneducated while only 14% manage to obtain education beyond primary schooling. Interestingly, sons who head split-offs are those who also experience greater downward mobility in schooling. This, in part, explains why spit-offs tend to experience greater downward mobility in wealth.

³³ This is because of the fact that new households are mostly set-up in the same bari.

³⁴ The estimate for the USA data is 0.37, before the transfer of bequests (Charles and Hurst, 2003).

A similar persistence is observed in occupation, particularly among the sons of self-employed individuals. However, substantial movement out of agricultural labour is observed: the probability that sons of agricultural wage labourers move into waged work and self-employment in non-farm sector is 29% and 36% respectively.

The above findings have serious implications for the process of economic development. Large persistence over the life-cycle implies that poverty-trap may exist: individuals who continue with poorer wealth may continue to remain poor over a longer period of time. For many of these individuals, this is unlikely to change across generations, as evidenced in the lack of intergenerational mobility in wealth. One policy option to remove these disadvantages in the initial condition is to equalize educational opportunities.

The lack of educational mobility in Matlab villages is striking. And it remains the most important determinant of economic mobility in our data. Comparisons of intergenerational schooling mobility and the elasticity of son's schooling with respect to father's wealth between two cohorts suggest that schooling persistence has not declined much over time. Large scale investment in rural schooling infrastructure currently undertaken by the government is therefore well-placed. The success of these investments in raising economic mobility in later life would, nevertheless, depend on the progressivity of these investments i.e. to what extent they facilitate increase schooling attainment for all groups, particularly the children of uneducated and poor parents.

Lastly, we find intergenerational and temporal mobility to be of similar magnitude in our data. To the extent ability is more correlated between periods of a person's life than generations of the same family, our data suggests that ability plays a smaller role in determining economic mobility in the rural area. Whilst plausible, for a lack of data, we are unable to distinguish between other competing hypothesis of persistence such as credit constraints and preference heterogeneity. Future studies should explore the exact cause behind intergenerational persistence in economic status.

References

- Angrist, Joshua D. and William Evans (1998). "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size", *American Economic Review*, vol. 88(3), pp 450-77.
- Andrade, Eduardo, Fernando Veloso, Regina Madalozzo, and Sergio G. Ferreira (2002). "Do Borrowing Constraint Decrease Intergenerational Mobility? Evidence from Brazil" (mimeo).
- Asadullah, M. Niaz (2006). "Returns to education in Bangladesh". *Education Economics* (forthcoming).
- Atkinson, A. B., François Bourguignon and Christian Morrisson (1992). *Empirical Studies on Earnings Mobility*. Chur, Philadelphia: Harwood Academic Publishers.
- Banerjee, Abhijit and Andrew F. Newman, (1993). "Occupational Choice and the Process of Development", *Journal of Political Economy*, vol. 101(2).
- Berman, Eli and Stepanyan, Ara (2003). "Fertility and Education in Radical Islamic Sects: Evidence from Asia and Africa", NBER discussion paper.
- Bourguignon, Francois, Francisco Ferreira and Marta Menedez (2003). "Inequality of Outcomes and Inequality of Opportunities", William Davidson Institute Working Paper no. 630.
- Charles, Kerwin and Erik Hurst (2003). "The Correlation of Wealth Across Generations", *Journal of Political Economy*, 111(6), pp. 1155-1182.
- Drèze, Jean, Peter Lanjouw and Naresh Sharma (1998). "Economic Development in Palanpur 1957-93". In: Peter Lanjouw and Nick Stern, ed. *Economic development in Palanpur over five decades*. Oxford University Press.
- Dunn, Christopher (2003). "Intergenerational Earnings Mobility in Brazil and its Determinants", University of Michigan (mimeo).
- Emran, M. Shahe, Otsuka, Misuzu and Shilpi, Forhad, (2003). "Gender, Generations, and Non-Farm Participation", World Bank discussion paper.
- Fields, Garry (2000). *Distribution and Development*. New York: Russell Sage Foundation.
- Filmer, Deon and Lant Pritchett (2001). "Estimating Wealth Effect without Expenditure or Income data", *Demography*.
- Foster, Andrew (1993). "Household Partition in Rural Bangladesh", *Population Studies*.
- Foster, Andrew and Mark Rosenzweig (2001). "Household Division and Rural Economic Growth", *Review of Economic Studies*, vol. 69(4).
- Fuwa, N. (1999). "An Analysis of Social Mobility in a Village Community: The Case of a Philippine Village", *Journal of Policy Modelling*, vol. 21(1).
- Grawe, Nathan and Casey B. Mulligan (2002). "Economic Interpretations of Intergenerational Correlations", *Journal of Economic Perspectives*, vol. 6(3).
- Grawe, Nathan (2004). "Intergenerational Mobility for Whom? The Experience of High- and Low-Earnings Sons in International Perspective." In Miles Corak, ed. *Generational Income Mobility in North America and Europe*. Cambridge: Cambridge University Press.
- Hossain, Mahabub, Manik Bose, Alamgir Chowdhury and Ruth Meinzen-Dick (2002). "Changes in Agrarian Relations and Livelihood in Rural Bangladesh: Insights from Repeat Village Studies". In: Ramachandran, V.K. and Swaminathan, Madhura ed. *Agrarian Studies: Essays on Agrarian Relations in Less Developed Countries*. New Delhi: Tulika Books.
- Mahmud, Wahiduddin (1996). "Employment Patterns and Income Formation in Rural Bangladesh: The Role of Rural Non-farm Sector", *Bangladesh Development Studies*, vol. 24(3).
- Rahman, Omar, Jane Manken, Andrew Foster, and Paul Gertler (2001). "Matlab Health

- and Socioeconomic Survey (MHSS), 1996: Overview and User's Guide". ICPSR, Ann Arbor, Michigan.
- Razzaque, Abdur, Lutfun Nahar, A. M. Sarder, J. Ginneken and M.A. Shaikh (1998). *Demographic Surveillance System-Matlab*, Scientific report no 83., ICDDR,B.
- Rosenzweig, Mark (2003). "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility", *American Economic Review* papers and proceedings.
- Sahn, David E. and, David Stifel (2003). "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", *Review of Income and Wealth*, vol. 49(4).
- Sen, Binayak (1996). "Rural Non-farm Sector in Bangladesh: Stagnating and Residual, or Dynamic and Potential?", *Bangladesh Development Studies*, vol. 24(3).
- _____ (2003). "Drivers of Escape and Descent: Changing Household Fortunes in Rural Bangladesh", *World Development*, vol. 31(3).
- Swaminathan, Madhura (1991). "Measuring Mobility in Wealth: Estimates from a South Indian Village", *Oxford Bulletin of Economics & Statistics*, vol. 53(2).
- Solon, Gary (2004), "A Model of Intergenerational Mobility Variation over Time and Place", In Miles Corak, ed. *Generational Income Mobility in North America and Europe*. Cambridge: Cambridge University Press.
- _____ (2002). "Cross-Country Differences in Intergenerational Earnings Mobility", *Journal of Economic Perspective*, vol. 16(3).
- _____ (1992). "Intergenerational Income Mobility in the United States", *American Economic Review*, vol. 82(3).
- Valentino, Dardanoni (1993). "Measuring Social Mobility", *Journal of Economic Theory*, vol. 61.
- Walker, Thomas and J. Ryan (1990). *Village and Household Economics in India's Semi-arid Tropics*. Baltimore: The Johns Hopkins University Press.

Appendix

Appendix Table 1: School availability (by types) in Matlab villages, 1996

School Types	Primary level	Secondary level
NG Madrassas	6	2
G Schools	63	0
NG Schools	11	18
Sub-total	80	20

Appendix Table 1a: Distribution of household landholding in Matlab villages, 1996

	Cultivable land		Homestead land	
	Freq.	Percent	Freq.	Percent
Landless	15,219	38.14	1,263	3.17
Less than .20 acre	5,373	13.47	33,357	83.6
.21 to .40	4,826	12.09	3,669	9.19
.41 to 1.00	7,633	19.13	1,424	3.57
1.01 to 2.00	4,218	10.57	169	0.42
2.01 and above	2,633	6.6	21	0.05
Total	39,902	100	39,902	100

Appendix Table 1b: Distribution of households by missing retrospective data and sex of the head

Data missing	Head is male	Head is female	Total
Yes	1585 (13.2)	632 (22.15)	2217
No	10430 (86.8)	2221 (77.85)	12651
N	12015	2853	14868

Note: Column percentage in parenthesis.

Appendix Table 1c: Mean statistics of sample of sons

Variable	Full		Inherit		Split-off	
	Mean	SD	Mean	SD	Mean	SD
Age	39.00	8.72	41.33	9.39	36.42	7.07
Age_hsq	1597.10	745.03	1796.90	833.59	1376.22	554.82
Non muslim	0.15	0.36	0.16	0.37	0.14	0.35
Married	0.96	0.20	0.94	0.23	0.97	0.16
Household size	5.50	2.00	6.00	2.22	4.94	1.55
Spousal education	1.98	2.84	2.01	2.89	1.94	2.78
Schooling (numbers of grade completed)	3.22	3.70	3.59	3.82	2.82	3.52
Self-employed in agriculture	0.30	0.46	0.34	0.47	0.26	0.44
Self-employed in non-agriculture	0.35	0.48	0.33	0.47	0.38	0.49
Wage-employed in non-agriculture	0.23	0.42	0.23	0.42	0.24	0.43
Wage-employed in (agricultural) labour	0.07	0.25	0.06	0.24	0.08	0.27
Unemployed	0.01	0.11	0.02	0.13	0.01	0.08
Others	0.00	0.06	0.00	0.06	0.00	0.05
OUemp (others + unemployed)	0.02	0.13	0.02	0.15	0.01	0.10
First child is a daughter	0.40	0.49	0.38	0.48	0.43	0.50
Wealth index	1.64	0.66	1.76	0.67	1.52	0.64
Alternative wealth index (1)	-0.26	1.58	-0.02	1.60	-0.53	1.50
Alternative wealth index (2)	0.39	0.23	0.43	0.24	0.35	0.22
Residence in treatment area	0.56	0.50	0.56	0.50	0.57	0.50
Inherited household	0.53	0.50	-	-	-	-
Birth order in 1974	1.78	0.98	-	-	-	-
Gap between own schooling and maximum of that among siblings	1.38	2.64	-	-	-	-
Parental characteristics (in 1974)						
Father's age	54.72	10.79				
Father's wealth index	1.14	0.66	1.17	0.66	1.10	0.66
Father's alternative wealth index (1)	-0.06	1.59	0.01	1.61	-0.14	1.58
Father's alternative wealth index (2)	0.27	0.19	0.28	0.19	0.26	0.19
Father's education	1.94	2.78	2.18	2.94	1.67	2.56
Mother's education	0.36	1.18	0.42	1.30	0.29	1.04
Mother's education missing	0.08	0.26	0.09	0.28	0.06	0.24
Father self-employed in agriculture	0.47	0.50	0.48	0.50	0.46	0.50
Father self-employed in non-agriculture	0.16	0.37	0.15	0.35	0.18	0.38
Father wage-employed in non-labour	0.10	0.30	0.11	0.31	0.10	0.30
Father wage-employed in (agricultural) labour	0.20	0.40	0.19	0.39	0.22	0.41
Father unemployed	0.05	0.21	0.06	0.23	0.03	0.18
Father in other jobs	0.02	0.12	0.02	0.13	0.01	0.12
N	5044		2648		2396	

Note: Wealth index used weights from coefficients on assets in a regression of household asset value (in logs). Alternate wealth indices, 1 and 2, are constructed following principal component analysis and weights derived from household expenditure regression respectively.

Appendix Table 1d: Selected attributes of households by missing retrospective data

Variable	Sample with missing retrospective data		Sample with non-missing retrospective data	
	Mean	SD	Mean	SD
Head's age	47.16	12.65	47.72	14.06
Head is Hindu	0.10	0.30	0.14	0.35
Head's is married	0.95	0.21	0.95	0.21
Household size	5.45	2.06	5.70	2.22
Head's schooling	4.07	4.07	3.00	3.58
Household wealth index	1.76	0.71	1.74	0.67
Alternative wealth index (1)	0.20	1.99	-0.04	1.59
Alternative wealth index (2)	0.44	0.27	0.42	0.24
Total homestead land (in acres)	11.77	18.80	11.59	16.27
Total cultivable land (in acres)	51.61	114.23	65.01	111.36
Residence in treatment area	0.52	0.50	0.56	0.50
Number of households in bari	9.35	8.58	10.53	8.11
No other household in bari	0.07	0.25	0.03	0.18
N	1585		10430	

Appendix Table 2: Summary statistics and (regression) weights relating to household asset and housing quality variables used in the construction of household wealth index

	(1)	(2)	MHSS, 96		MSEC, 96	
	Asset	LnPCE	mean	Stan.dev	mean	Stan.dev
Household has a cow	0.191** (4.97)**	0.041 (2.10)*	0.42	0.49	0.33	0.47
Household has a boat	0.049 (1.19)	0.079 (3.76)**	0.31	0.46	0.28	0.45
Household has a radio	0.311 (6.93)**	0.103 (4.47)**	0.48	0.50	0.39	0.49
Household has a watch	0.129 (2.72)**	0.156 (6.41)**	0.59	0.49	0.49	0.50
Household has a hurricane	0.106 (1.59)	-0.009 (0.25)	0.91	0.29	0.88	0.33
Household has a quilt	0.348 (7.23)**	0.127 (5.13)**	0.62	0.48	0.54	0.50
Roof of largest room made of tin	0.535 (4.52)**	0.072 (1.18)	0.97	0.16	0.96	0.19
Wall of largest room made of tin	0.656 (14.82)**	0.160 (7.04)**	0.52	0.50	0.43	0.50
Tube well (source of drinking water)	0.401 (4.75)**	0.023 (0.53)	0.95	0.22	0.94	0.24
Tube well (source of cooking water)	0.221 (1.90)+	0.057 (0.95)	0.06	0.24	0.04	0.20
Tube well (source of bath water)	0.171 (1.41)	0.142 (2.29)*	0.06	0.24	0.03	0.18
Tube well (source of water for washing)	0.274 (1.75)+	0.221 (2.76)**	0.03	0.17	0.01	0.12
Adjusted R ²	0.38	0.28				
N	3423	3421	3423		5044	
Mean of dependent variable	10.64	9.22				
Stan. Dev. of dependent variable	1.35	0.64				

Note: Absolute value of t-statistics in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%. All the variables apart from the dependent variables are dummies. Column (1) refers to the regression where the dependent variable is household asset value. It also controls for education of the head, his spouse and household size. Column 2 (regression of household expenditure) controls for education of the head and his spouse. Both the regressions use data for male-headed households only.

Appendix Table 3: Regression estimates of intergenerational correlation in wealth [Dependent variable: son's wealth]

	1	2	3	4	5	6	7	8	9	10	11	12	13
	OLS	OLS	IV	IV	OLS	IV	IV	FE	FE-IV	FE-IV	FE	FE-IV	FE-IV
Father's wealth	0.538 (50.14)**	0.357 (28.21)**	0.546 (9.06)**	0.476 (6.57)**	0.37 (29.03)**	0.444 (6.03)**	0.492 (8.32)**	0.303 (15.30)**	0.49 (4.12)**	0.561 (6.47)**	0.292 (14.78)**	0.457 (3.73)**	0.535 (6.09)**
Schooling		0.047 (20.48)**	0.034 (7.13)**	0.041 (6.99)**	0.048 (20.75)**	0.042 (7.09)**	0.038 (7.48)**	0.042 (14.71)**	0.037 (5.81)**	0.032 (6.86)**	0.033 (10.82)**	0.03 (4.79)**	0.025 (5.33)**
Household size		0.049 (12.98)**	0.049 (12.49)**	-0.083 (2.24)*	0.046 (12.07)**	-0.101 (2.64)**	-0.082 (2.42)*	0.057 (11.89)**	-0.141 -1.63	-0.06 -0.94	0.057 (11.99)**	-0.154 (1.75)+	-0.068 -1.08
Self-employed, non-agri.		-0.066 (3.44)**	-0.037 (1.67)+	-0.068 (2.59)**	-0.021 -1.18	-0.068 (2.58)**	-0.057 (2.29)*	-0.017 -0.73	-0.038 -1.24	-0.033 -1.2	-0.021 -0.91	-0.045 -1.41	-0.039 -1.4
Wage-employed, non-agri.		-0.086 (4.33)**	-0.055 (2.43)*	-0.106 (3.52)**	-0.046 (2.42)*	-0.109 (3.60)**	-0.094 (3.39)**	-0.024 -1	-0.066 (1.97)*	-0.049 (1.72)+	-0.03 -1.26	-0.076 (2.21)*	-0.058 (2.01)*
Wage-employed, agri. labour		-0.221 (7.50)**	-0.189 (5.97)**	-0.282 (6.26)**	-0.218 (6.99)**	-0.289 (6.33)**	-0.268 (6.39)**	-0.205 (5.41)**	-0.292 (4.27)**	-0.238 (4.40)**	-0.211 (5.60)**	-0.307 (4.39)**	-0.249 (4.59)**
OUemp		-0.109 (1.90)+	-0.084 -1.42	-0.139 (1.89)+	-0.083 -1.48	-0.165 (2.20)*	-0.154 (2.17)*	-0.077 -1.2	-0.168 (1.98)*	-0.142 (1.91)+	-0.078 -1.22	-0.175 (2.03)*	-0.147 (1.99)*
Inherited household					0.095 (6.04)**	0.207 (6.79)**	0.227 (4.28)**	0.049 (2.81)**	0.171 (3.32)**	0.175 (2.29)*	0.052 (3.03)**	0.179 (3.42)**	0.185 (2.42)*
Spousal education											0.026 (7.48)**	0.025 (5.21)**	0.024 (5.91)**
Constant	1.031 (71.20)**	0.882 (7.63)**	0.769 (6.15)**	0.433 (2.70)**	0.795 (6.90)**	0.372 (2.27)*	0.397 (2.49)*	1.224 (8.60)**	0.437 -1.25	0.645 (2.25)*	1.226 (8.69)**	0.42 -1.18	0.636 (2.21)*
Adjusted R ²	0.29	0.39			0.39			0.25			0.26		
Bari fixed effects	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Ovrid			0.35	0.04		0.25	0.28		0.09	0.01		0.13	0.01
Exogeneity test			0	0		0	0						
N	5044	5044	5044	5044	5044	5044	5044	5044	5044	5044	5044	5044	5044

Note: Robust t-stats are reported. + significant at 10%; * significant at 5%; ** significant at 1%. All regressions control for individual's age, age-squared and religion. Model (3) instruments father's wealth (in 1974) by father's occupation and education. Models 4 and 6 additionally instrument son's family size by "treatment area dummy" and "sex of eldest child" (of the son). In addition to father's wealth and son's family size, Model 7 also instruments household inheritance by i) head's birth order (in 1974) and (ii) difference between head's schooling and the maximum of that among his siblings. Model 8-13 repeats models (5)-(7) with bari fixed effects (1775 dummies) and with and without control for spousal education. Test of exogeneity (of instrumented variables i.e. father's wealth, son's household size and so on) is based on Durbin-Wu-Hausman test with the null that the variable is exogenous. Over-identification test (overid) is based on Hansen's J-statistics. Only p-values are reported for the two tests.

Appendix Table 4: Regression estimates of intergenerational correlation in wealth, inherited and split-off households [Dependent variable: son's wealth]

	Inherited				Split-offs			
	OLS	OLS	IV	IV	OLS	OLS	IV	IV
Father's wealth	0.549 (37.99)**	0.382 (21.34)**	0.643 (8.06)**	0.516 (4.13)**	0.507 (31.98)**	0.339 (18.92)**	0.368 (4.11)**	0.354 (3.80)**
Age_h		-0.019 (2.49)*	-0.017 (2.20)*	0.022 -1.05		0.001 -0.08	0.001 -0.12	0.039 (2.07)*
Age_hsq		0 (2.94)**	0 (1.88)+	0 -0.95		0 -0.96	0 -0.83	0 -1.06
Hindu		-0.051 (1.70)+	-0.016 -0.48	-0.043 -1.13		-0.002 -0.06	0 -0.01	-0.041 -1.08
Schooling		0.045 (14.31)**	0.025 (3.83)**	0.038 (3.46)**		0.046 (13.17)**	0.044 (6.56)**	0.043 (6.17)**
Household size		0.046 (10.28)**	0.045 (9.60)**	-0.081 -1.32		0.036 (4.86)**	0.036 (4.75)**	-0.073 (1.69)+
Self-employed in non-agri.		-0.095 (3.61)**	-0.056 (1.87)+	-0.106 (2.37)*		-0.016 -0.56	-0.011 -0.35	-0.008 -0.24
Wage-employed in non-agri.		-0.089 (3.25)**	-0.05 -1.62	-0.111 (2.27)*		-0.064 (2.23)*	-0.059 (1.79)+	-0.079 (2.23)*
Wage-employed, agri. labour		-0.287 (6.51)**	-0.245 (5.19)**	-0.36 (4.40)**		-0.142 (3.63)**	-0.137 (3.28)**	-0.179 (3.85)**
OUemp		-0.1 (1.46)	-0.071 (1.00)	-0.104 (1.17)		-0.172 (1.90)+	-0.169 (1.85)+	-0.285 (2.80)**
Constant	1.111 (52.75)**	1.284 (8.12)**	1.1 (6.15)**	1.025 (5.42)**	0.963 (49.12)**	0.661 (3.24)**	0.638 (2.96)**	0.295 -1.11
Adjusted R ²	0.3	0.39			0.27	0.37		
Ovrid			0.27	0.07			0.99	0.99
Exogeneity test			0	0			0.74	0.02
N	2648	2648	2648	2648	2396	2396	2396	2396

Note: Excluded occupation dummy is self-employment in agriculture. First four columns refer to inherited-household sample whereas the last four are for split-offs. Instruments for father's wealth (in 1974) are father's occupation while instruments for son's family size are "treatment area dummy" and "sex of eldest child" (of the son). Robust standard errors are reported. Test of exogeneity (of father's wealth and son's household size) is based on Durbin-Wu-Hausman test with the null that the variable is exogenous. Over-identification test (overid) is based on Hansen's J-statistics. Only p-values are reported for the two tests.

Appendix Table 5: Wealth transition matrices, father-son sample

	Raw wealth data	Age-adjusted wealth data	Age & education adjusted wealth data	Age, education & family size adjusted wealth data	Age, education, family size & occupation adjusted wealth data
	Son	Son	Son	Son	Son
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Full sample					
	1 0.42 0.29 0.18 0.09 0.01	1 0.36 0.27 0.21 0.11 0.04	1 0.31 0.29 0.21 0.11 0.09	1 0.31 0.28 0.21 0.11 0.09	1 0.29 0.29 0.2 0.12 0.1
	2 0.28 0.26 0.24 0.15 0.07	2 0.28 0.25 0.2 0.14 0.13	2 0.25 0.25 0.2 0.15 0.16	2 0.25 0.25 0.2 0.15 0.16	2 0.24 0.25 0.2 0.14 0.17
(Father)	3 0.18 0.23 0.23 0.25 0.11	3 0.2 0.25 0.21 0.18 0.16	3 0.21 0.24 0.19 0.18 0.18	3 0.21 0.24 0.19 0.18 0.18	3 0.22 0.24 0.19 0.17 0.19
	4 0.1 0.15 0.21 0.35 0.19	4 0.12 0.14 0.22 0.25 0.27	4 0.14 0.14 0.19 0.24 0.29	4 0.14 0.14 0.19 0.24 0.29	4 0.15 0.14 0.21 0.23 0.27
	5 0.04 0.08 0.15 0.39 0.35	5 0.06 0.08 0.15 0.32 0.39	5 0.09 0.09 0.2 0.33 0.29	5 0.09 0.09 0.2 0.33 0.29	5 0.1 0.08 0.2 0.34 0.28
Pearson's χ^2		1353	970	573	569
Likelihood χ^2		1429	1026	598	594
Inherited					
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
	1 0.38 0.29 0.19 0.13 0.02	1 0.39 0.23 0.2 0.13 0.06	1 0.32 0.26 0.18 0.12 0.11	1 0.33 0.26 0.17 0.13 0.11	1 0.31 0.26 0.17 0.13 0.13
	2 0.25 0.26 0.22 0.18 0.1	2 0.29 0.22 0.17 0.16 0.16	2 0.26 0.21 0.19 0.17 0.19	2 0.25 0.21 0.18 0.17 0.19	2 0.24 0.22 0.19 0.15 0.19
(Father)	3 0.14 0.19 0.24 0.27 0.15	3 0.18 0.23 0.2 0.21 0.18	3 0.21 0.2 0.19 0.19 0.21	3 0.21 0.19 0.19 0.19 0.21	3 0.2 0.2 0.18 0.18 0.23
	4 0.07 0.1 0.18 0.4 0.24	4 0.11 0.1 0.22 0.27 0.3	4 0.14 0.11 0.17 0.24 0.33	4 0.14 0.11 0.18 0.24 0.33	4 0.15 0.11 0.19 0.27 0.29
	5 0.03 0.04 0.11 0.41 0.4	5 0.05 0.05 0.11 0.34 0.45	5 0.07 0.06 0.19 0.38 0.3	5 0.07 0.06 0.19 0.38 0.31	5 0.08 0.05 0.22 0.37 0.29
Pearson's χ^2		768	585	352	347
Likelihood χ^2		824	618	369	363
Split-offs					
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
	1 0.46 0.3 0.17 0.06 0.01	1 0.33 0.33 0.23 0.09 0.02	1 0.28 0.31 0.25 0.1 0.06	1 0.28 0.31 0.25 0.1 0.06	1 0.27 0.33 0.23 0.11 0.06
	2 0.31 0.26 0.26 0.12 0.05	2 0.27 0.28 0.23 0.11 0.11	2 0.24 0.29 0.22 0.13 0.13	2 0.24 0.29 0.22 0.13 0.13	2 0.24 0.28 0.2 0.13 0.14
(Father)	3 0.22 0.26 0.23 0.22 0.06	3 0.22 0.28 0.23 0.14 0.13	3 0.21 0.28 0.2 0.16 0.14	3 0.21 0.28 0.2 0.16 0.14	3 0.23 0.27 0.2 0.15 0.15
	4 0.14 0.21 0.25 0.28 0.13	4 0.12 0.18 0.23 0.23 0.23	4 0.14 0.18 0.21 0.23 0.24	4 0.14 0.17 0.21 0.23 0.24	4 0.15 0.18 0.24 0.2 0.24
	5 0.05 0.11 0.2 0.36 0.28	5 0.07 0.11 0.2 0.3 0.32	5 0.12 0.13 0.21 0.27 0.28	5 0.12 0.13 0.21 0.27 0.28	5 0.13 0.12 0.18 0.3 0.27
Pearson's χ^2		602	399	236	238
Likelihood χ^2		619	427	248	242

Note: Underlying mobility indices are reported in Table 4.