Household Recombination, Retrospective Evaluation, and the Effects of a Health and Family Planning Intervention

Andrew Foster and Sveta Milusheva

April 17, 2015

Abstract

Analysis of the long term effects of social and public health programs using household survey data requires an understanding of patterns of household recombination–that is the processes by which households divide and fuse over time. In this paper we examine, in particular, the effects on educational mobility of a well-known maternal and child health and family planning program in Matlab Bangladesh. Using a novel resampling procedure that relies on longitudinally collected demographic surveillance data, we correct for biases that arise from household recombination that occurred subsequent to a baseline census but prior to the collection of the first round of detailed economic data. Our results suggest that the program resulted in a small increase in consumption per capita, decreases in family sizes, small changes in household recombination, and increases in child schooling except among the lowest education households. We also show that approximate corrections for these biases using more limited data are reasonably effective.

1 Introduction

While the process of economic mobility is often studied by examining the long term prospects of individuals with different background characteristics, in practice changes in the well-being of individuals are importantly determined by the households and families in which these individuals are embedded. This relationship is perhaps most clear with respect to household level measures such as expenditure or poverty, but even arises in the study of mobility with respect to something as individualized as education. Consider the question, for example, of whether educational inequality at one point in time persists into subsequent periods and the degree to which this persistent inequality is influenced by various types of public programs and services. One cannot simply look at educational change for a given person over twenty years. Similarly one cannot learn much by comparing the education of a child of one mother to the education of another child born to the same mother twenty years later. Childbearing and education are intrinsically tied to the span of childbearing of the mother and to the particular ages of the children. Moreover educational decisions happen within the context of households and multiple household factors may be responsible for the processes of educational evolution over time. From this perspective it would seem advantageous to look at educational mobility from the stand point of a particular family line, comparing the education of children of a particular age at one point in time to the education of children of that same age some years later in that same line.

This approach might be relatively straightforward using long-term household panel survey data if households stayed fixed over time; however, they do not. A household at one point in time morphs through a process of household division and fusion, which we combine under the rubric of household recombination. Not only is it necessary to account for this process of recombination in the evaluation of data on economic mobility but the process of recombination in turn can affect the process of economic mobility. Depending on the nature of the sampling processes and how recombination proceeds over time it may be necessary to reweight data based on weights that may themselves be endogenous with respect to household recombination.

The importance of understanding household formation as an element of analyses of other development outcomes has been previously recognized in the literature. Foster 1993, for example, noted "It is increasingly recognized that certain demographic variables should be treated as endogenous in analyses of economic and demographic data from developing countries, very little is known about the implications of the fact that joint residence is itself a choice variable." In subsequent work looking at the relationship between household division and inequality, Foster and Rosenzweig 2002 argue "An improved understanding of the determinants of household division is thus useful not only for dealing with the potential selectivity of panel designs that drop dividing households, but in studying household behaviour and income change generally." But both papers focus only on the process of household division and, in part as a consequence of this, neither deals explicitly with the issue of how sampling composition is affected by the process of household recombination and thus influences the construction of sample weights.

The process of using weights to adjust for sample attrition has, of course, achieved substantial attention in the literature (Fitzgerald, Gottschalk, and Moffitt 1998, Moffit, Fitzgerald, and Gottschalk 1999) . Generally, one inflates the weight of observed sampling units households based on the assumption that attrition is random with respect to processes of interest in the data conditional on the observables. Of course, this assumption may not be correct in general and even if it is, the outcomes of the attritting population may be sufficiently distinct from the observed population (e.g., attrition through mortality) that one cannot sensibly combine the outcomes of the observed and attritted population into a single metric. But little attention has been given to the question of weighting in a setting in which the problem is not sample attrition but the process of household recombination. In this case observability may not be as much of an issue as in the case of attrition, but the flow of people across households and the fact that the sampling unit for a survey is generally the household creates a new set of problems.

To understand this point, consider a random sample of households collected at time t and assume that larger (at time t) households are more likely to divide. If all descendant households are followed, then one will correctly measure at time t+1 the distribution of t+1 attributes such as household size. However, if these t+1 households are used to retrospectively construct the mean household size at time t, then the estimate will be too large because large t households are overrepresented in the resulting sample of t+1 households. A similar bias would arise if household size at time t were estimated from a random sample of households in period t+1. A simple correction in each case would be to inverse weight each household by the number, if available, of co-descendant households with the same antecedent household. As we detail below, however, the situation is further complicated in the presence of household fusion. The set of all descendant households in t+1 generated from an initial sample in period t will no longer yield an unbiased estimate of household measures at time t+1 because t+1 households composed of members of multiple antecedent households will be more likely to be selected than would be the case if t+1 households were selected randomly. To correctly constitute a represented sample one would need to know the sampling probabilities and descent paths of households that were not sampled at t.

This is not just a curiosity. Surveys are at times used to evaluate the consequences of interventions that were introduced at a previous period, and in some cases retrospective or previously collected data are incorporated into the analysis in order to estimate differential change over time. In such cases we may ask if it is possible to mimic the results of a randomized trial in which a baseline is collected from a random sample at a particular point in time, a set of treatments is assigned to the participants, and then outcomes are evaluated at some endline. Our answer is a tentative yes, but as our application suggests, the data requirements for doing so are extremely demanding.

In this paper, we tackle these larger questions of household recombination, sample selection, and weighting mechanisms in panel datasets by looking at a specific example. We focus, in particular, on the Matlab Health and Socio-economic Survey done in Matlab, Bangladesh in 1996, which can be linked to an ongoing process of population-wide vital registration and periodic censuses known as the Health and Demographic Surveillance System (HDSS) since 1966. There are dozens of papers and dissertations using this data, many of which use it to examine the effects of a Maternal and Child Health and Family Planning Program (MCH/FP) begun in 1978. This combination of an HDSS with an experimentally introduced intervention has been used in various other regions, including Navrongo (Ghana), Rakai (Uganda) and Filabavi (Vietnam) among others. The comprehensive data available through surveillance populations opens up a variety of possibilities for studying the long-run effects of the interventions, but it is not always clear how to make the link between survey sample and the vital registration data. Our paper demonstrates how in the case of an RCT or other intervention, it is possible to make use of the long-run HDSS data by creating weights that are endogenous to the recombination process but correctly assess population impacts.

There are two specific issues with the Matlab data sets that reflect our broader concerns about representativeness in panel data. First, the sampling frame for the

1996 survey data reflects patterns of residence, inclusive of those affected by the 1978 intervention, leading to a problem of endogenous sampling (Hausman and Wise 1981). The set of 1974 antecedents of this population is thus not necessarily representative of the pre-intervention population, potentially biasing estimates of the intervention. Second, the 1996 sample does not constitute a random set of the descendants of these 1974 antecedents, thus complicating estimates of mobility based on a 1974 frame. We correct for both the non-representative nature of the antecedents of the 1996 sample and the incomplete sampling of descendants of these antecedents by devising appropriate weights. The weights are constructed using a unique resampling procedure that is possible due to the nature of HDSS data. Results from our optimal procedures are then compared to the results of related procedures that can be carried out using more limited data.

The effect of the program on assets and income have been looked at by Joshi and Schultz in several of their papers, which analyze a number of long-term effects of the MCH/FP intervention (Joshi and Schultz 2007, Schultz 2009). Joshi and Schultz 2007 analyze women's income, but also look at household assets. They find that all but non agricultural assets are significantly larger in treatment areas than comparison areas, but that gains are generally concentrated among better educated older women. They describe this as being consistent with physical assets being a substitute for a decrease in children for better educated women. Asadullah 2012 looks specifically at household wealth mobility using the Matlab data and focusing on the relationship between fathers' and sons' wealth as well as changes in ones own wealth between 1974 and 1996 for men who were a head of household at both times. Asadullah finds limited intergenerational wealth mobility and suggests that this might be due to limited schooling mobility. But Asadullah does not consider the sampling bias and household formation and recombination. If there is bias in the selection of the sample so that those selected are not representative of the population, then the treatment effect calculated is only for that selected group, which might experience a differential impact from the program. In addition, household recombination leads to a sample of descendants that is not representative because the changes in household structure could be correlated with characteristics of interest such as income in 1996, leading to a correlation with the probability of being in the sample.

We address a variety of issues related to economic mobility including changes in family size, the number of descendant households and consumption growth. But we give specific attention to changes in educational investment. Education is an outcome of primary interest for several reasons. The idea that child education can be importantly influenced by reductions in the cost of fertility control is thought to be an important motivation for a public subsidy of fertility control given possible externalities to child schooling and the consequent public subsidies to child schooling. Indeed, there is an extensive literature testing for a tradeoff between the quantity and quality that has bearings on the relative substitutability of these two outcomes from the perspective of parents (Rosenzweig and Wolpin 1980, Becker and Lewis 1974). Moreover, given the relatively long time frames between fertility reduction and the realization of educational achievement this approach does not lend itself well to the kinds of prospectively studied randomized interventions that are frequently used to study more short term outcomes.

This question has been examined in Matlab with 1982 and 1974 census data using dynamic decision rules by Roy and Foster 1996. This approach focuses on short run change by looking at the same mother at two points in time. The question has also been examined cross-sectionally in the Matlab population based on the 1996 data (Joshi and Schultz 2007). We argue that this latter approach is not definitive for two reasons. First, this cross-sectional approach does not allow one to account for differences between the treatment and comparison areas that were in place prior to the introduction of the program. By conditioning on earlier differences, one can control at least in part for the possibility that these early differences, rather than the treatment program itself, were responsible for any later observed differences. Second, the cross-sectional approach does not allow one to condition on pre-intervention educational attainment and thus to look at the consequences of the program for educational mobility.

We focus specifically on the effect of the program on education of children 6-16, controlling for a set of characteristics of antecedant 1974 households including predicted 1974 consumption, 1974 family size, and 1974 child's education. We find that those households in the bottom third of the 1974 education distribution did not benefit from the treatment in terms of educational attainment but those at the top two thirds did. In addition we find that the program led to a significant decrease in family size, as might be expected, a small decrease in the number of descendant households, and a small increase in consumption within certain strata. Our findings underscore the importance, when conducting longitudinal analysis, of considering household recombination and how this process influences sampling.

2 Data

The International Centre for Diarrhoeal Disease Research (ICDDR,B) in the Matlab region of Bangladesh began to maintain a Health and Demographic Surveillance System registering all births, deaths and migrations starting in 1966. There are data available for the full period on 149 villages, which include over 200,000 people. In 1974 the ICDDR,B conducted the first comprehensive census of the region. Censuses were again conducted in 1982, 1993 and 1996. From this census data we have information on every single household in the region including basic demographic information and some information on assets.

Along with collecting detailed demographic data, the ICDDR,B initiated a Maternal and Child Health and Family Planning Program (MCH/FP) intervention in 1977. This intervention was implemented in 70 of the villages, and as part of it, women of childbearing age received doorstep delivery of contraceptives and antenatal care, children received in-home vaccinations delivery, and there were increased services for the prevention and management of childhood diarrheal and acute respiratory illnesses (Fauveau 1994, Phillips et al. 1982). The family planning and maternal health portion of the intervention was first rolled out in 1977. The services were then expanded in 1982 to include vaccinating children under 5 for measles, DPT, polio and tuberculosis, as well as providing vitamin A supplementation. Such services were not provided by the government to the rest of the region until 1988.

For this intervention, contiguous villages were grouped together into treatment and control areas, partly to control for contamination and spillover effects between treated and not treated villages. Nevertheless, it is a case-control study where treatment and control areas were reasonably matched on observables, though the matching was not perfect. Much of the research on Matlab using ICDDR,B data focuses on the outcomes of this intervention.

In addition to the censuses that were collected, in 1996 a Health and Socio-economic Survey (MHSS) was conducted in Matlab. This survey collected detailed economic and social data on a sample of the population, which had not been done before. The goal was to use this data in order to look at the effect of the MCH/FP intervention on a wide range of outcomes and over a long period (Rahman et al. 1999).

The MHSS sample was selected from the population in 1993. In 1993, there were 38,489 households split among 7,440 baris, clusters of households in close physical proximity that are usually linked in a kin network. Of those, 2,883 baris were randomly picked to be part of the sample. Within each bari, one household was randomly selected for the detailed interview. Sampling was done at the bari level because it provides a better representation of family networks as compared to sampling households. There were 102 baris that no longer existed in 1996, and therefore the final number of baris sampled was 2,781. A second household in each bari was also interviewed, but this was not done randomly, so most researchers conduct analyses using only the first household, and we also focus on this primary sample.

The detailed level of the data collected by the MHSS on income, consumption, education, health, and other outcomes has allowed researchers to examine the long run impact of the 1978 MCH/FP program. There are dozens of papers and dissertations that use the data to look at various outcomes (Joshi and Schultz 2013, Field and Ambrus 2008, Maitra 2003, Barham 2012). In order to calculate the effect of the treatment on the whole population, these papers use the sampling weights included in the dataset. As we will describe in the following subsection, certain aspects of the sampling procedure could bias the results depending on the questions of interest. In particular, it is not necessarily appropriate to use the weights given in the dataset when looking at educational mobility.

3 Creating Representative Weights for 1974

As noted, the specific issues that must be addressed in these data arise from the fact that households were picked based on the 1993 bari configurations and not based on the pre-treatment 1974 population. Because households were picked based on 1993 configurations, the inverse probability weights available in the dataset make the sample representative of the 1996 population.[1] But in order to

---

[1] In general households remain stable over the period 1993-1996;therefore, although the sample was picked based on the 1993 configurations, it is nevertheless representative of the 1996 popula-

be able to evaluate the results from an experiment, or in this case the effectiveness of the 1978 program on different groups, one wants the antecedents of the sample picked post intervention to be representative of the population before the intervention. A potential further complication is the sampling of households in 1996 may be in part a consequence of changes in household recombination resulting from the intervention.

The general problem is illustrated in Figure 1. The top row of the figure denotes a set of period t households divided into two baris. The bottom row shows the households at time t+1 with the arrows connecting the two rows showing the movement of individuals over time from period t to period t+1 households. Suppose the sampling strategy is to sample one of the two strata with probability 1/2 each and then within the strata to select one household. Then the probability of sampling, for example, the household in the upper left corner is 1/6. These probabilities in combination with the household descent mapping creates a set of probabilities that each of the period t+1 households appear in the sample. These probabilites are noted on the diagram. Now imagine that one's sample of t+1 households consists of all the descendants from the one household picked by the sampling scheme. Then the weights needed to create an unbiased estimate of some t+1 characteristic are dependent not only on the sampling probability of the particular household picked at period t but of any household that could have been picked in period t with a link to the t+1 household. The same is true in reverse if the t+1 sample is representative and one constructs a period t sample from the set of antecedents of the t+1 sample.

This latter issue is exactly the one that must be addressed in the context of the MHSS. Formally, define

- I_t is the set of households i at time t
- I_{tx} is the set of households i at time t of type x
- $A : I_{t+1} \Rightarrow I_t$ where $A(K)$ is the set of households in I_t that contain antecedents of members of household $K \subset I_{t+1}$
- J_{t+1x} is the set of households j at time $t+1$ such that $A(j) \subset I_{tx}$
- N_{tx} is the number of households i at time t of type x

---

tion.

- S_t is the sample drawn at time t

Then for some outcome c_{it} among a population with characteristics x, if

$$\bar{c}_{tN} = \frac{1}{N_{tx}} \sum_{i \in I_{tx}} c_{it} \tag{1}$$

and

$$\hat{c}_{tN} = \frac{1}{N_{tx}} \sum_{i \in I_{tx}} c_{it} \frac{\mathbb{1}(i \in A(S_{t+1}))}{\mathbb{E}(\mathbb{1}(i \in A(S_{t+1})))} \tag{2}$$

$$\hat{i}_{tN} = \frac{1}{N_{tx}} \sum_{i \in I_{tx}} \frac{\mathbb{1}(i \in A(S_{t+1}))}{\mathbb{E}(\mathbb{1}(i \in A(S_{t+1})))} \tag{3}$$

it may be shown that $\operatorname*{plim}_{N\to\infty} \frac{\hat{c}_{tN}}{\hat{i}_{tN}} - \bar{c}_{tN} = 0$. The ratio $\frac{\hat{c}_{tN}}{\hat{i}_{tN}}$ is simply the weighted average of the sampled c_{it} weighted using as weights the inverted probabilities of appearing in the antecedent sample.[2]

In short, in the Matlab context, in order for the sample of antecedent households to be representative of the population pre-treatment, it would have been necessary to randomly select baris from the 1974 population. Then one could, for example, pick one household per bari, trace which households descended from that household in 1996 and randomly pick one or more descendants per 1974 household. Instead, by randomly selecting descendants from the 1996 population for the 1996 sample without taking into account how many of them came from each 1974 household, the MHSS team inadvertantly exposed the sample to the outlined sources of bias. [3] By sampling baris rather than households, the bias could be mitigated because baris tend to be made up of households that are linked by kinship. Yet, women are likely to join the bari of their husband upon marriage, so a 1974 household with several daughters would have descendants in several baris. In addition, the

---

[2] Note that equation has a probability limit of one. We incorporate it in our formulation to make the analogy to a weighted sample (in which the sum of weights is normalized to one) and because dividing by the sample estimate leads to greater efficiency. Simply put if one's sample happens to be larger than its expected size then equation will also be larger than its expected size.

[3] One of the coauthors was on the original MHSS team and now recognizes the issues with the way the sampling was done, but at the time, the focus was on getting a representation of kin networks, which were assumed to be manifested in the bari structure, without considering the endogeneity of how kin networks might spread to other baris due to the formation and recombination of households.

decision of some descendants to split and form their own bari or to join a different bari could also be dependent on the observable and unobservable characteristics of the 1974 household. This dynamic would again affect the probability of a 1974 household being represented in the sample.

Of course, the extent to which the actual t sampling probabilities will differ from those that might be calculated on the basis of the t+1 sample population will depend on the extent to which period t households have multiple antecedents. To explore this issue we first must define the rules of descent. In particular for the purpose of this paper we define a 1974 household and a 1996 household to be linked if they meet one of three conditions:

1. Someone in the 1996 household is also in the 1974 household (Zero Order Link)
2. Someone in the 1996 household has lived with a member of the 1974 household at any point between 1974 and 1996 (First Order Link)
3. Someone in the 1996 household has lived with a person who lived with a member of the 1974 household prior to living with the 1996 person (Second Order Link)

These rules create a complex web of interrelationships. For illustrative purposes we randomly selected one 1996 household and found its antecedents. We then took the antecedent households and constructed each of their descendants. We then repeated this process. The resulting matrix of households is presented in Table 1. Each 1974 household has multiple descendant households and each 1996 household has multiple antecedent households. Indeed looking across the 23,913 households in the 1974 census that are linked to at least one 1993 household, we find that households have on average 3.62 (sd 3.04) descendant households living in the Matlab HDSS area but only .254 (sd. .529) descendant households that were part of the sample. Conversely the 34,365 households in the 1996 census have on average 2.35 (s.d. 1.66) descendants. It seems likely given the variation in the number of descendant households that the actual sampling probabilities for different 1974 households will be quite different from the sampling probabilities based on the 1996 sample.

The pattern of links is also relevant to the assessment of the intervention. Above we noted that estimates of the 1974 status could be conditioned on characteristics at time t, inclusive of whether a household lives in an area that would eventually receive the treatment. However, the 1996 outcomes for a particular household living in a treated area in 1974, given recombination, may be composed in part of indiiduals who were not in the treatment area. While this pattern would in a sense be of the flavor of "intent to treat", the effects of a treatment might in general depend on the treatment status of potential household partners. For example, Arunachalam and Naidu 2008 showed that dowries rose differentially in the treatment area. These higher dowries may have arisen because households in the study area preferred to place their daughters of marriageable age in the treatment area because of the better MCH services there. In such circumstances one might see a wealth increase in the treatment area, but the extent of this would depend on the fraction of treated households among possible marital partners. Fortunately, however, it seems that the process of descent in Matlab was sufficiently local that the vast majority of descendants living in a particular 1996 household had the same treatment status in 1974. Figure 2 presents the distribution of the average 1974 status of the individuals living in a particular 1996 household. As is evident a very high share of the households are mostly treatment or mostly comparison. Thus, it is reasonable to see the treatment effect as the effect of being in the treatment area when potential descendants have the same status.

The basic method used to resolve the problem of different sampling probabilities is to create a representative sample of 1974 households linked to the 1996 sample and assign weights based on the probability that a particular 1974 household is represented in the sample given the actual sampling rules. To find these probabilities, 1974 households are linked to all 1993 households. This is possible because there is data on the full population of 1974 and 1993 households, and the Demographic Surveillance System uses a unique ID that allows individuals to be traced at every point in time between 1974 and 1993. We thus know who they were living with at each point in time.

In this context, an antecedent is considered a 1974 household and a descendant is a 1996 household that is sampled based on the 1993 census. These antecedent-descendant links can be used to calculate the probability that a particular 1974 antecedent household has a descendant that appears in the 1996 sample. In the case of a simple random sample this calculation is straightforward. The probability that any particular 1974 household is picked is just $(1-(1-p)^n)$ if a 1974 household has n descendant households and the probability that any particular descendant household is picked is p. In the present case, however, because of the bari level sampling, the probability that particular households are picked is negatively correlated within baris with the extent of this correlation depending on bari size. Due to this complication in sampling, we calculate the probabilities by replicating the procedure implemented in 1996 to pick the sample based on the 1993 census. Note that this procedure would work for any arbitrarily complex sampling procedure and descent definition.

To elaborate, we took the 1993 population and randomly picked 2,883 baris from the total 7,440 baris, and then picked one household at random from each bari. This creates a sample of 2,883 households that consist of alternate MHSS 1996 samples. We did this 100,000 times. The antecedent-descendant links were used to establish which 1974 households were represented by at least one descendant household in each sample. The probability of a 1974 household being represented in an arbitrary 1996 sample, including the particular one observed, is the number of samples in which the household has at least one descendant out of 100,000 possible samples. We created probability weights by taking the inverse of the calculated probability and assigning that as the weight to each 1974 household.[4]

Having assigned each 1974 household a weight, we created a sample of 1974 households that is linked to the actual 1996 sample by taking all of the antecedents of the 1996 sample and grouping them together into what we are calling the "1974 sample." This consists of 5,319 households that all have at least one descendant in the actual 1996 sample. Using the 1974 probability weights calculated earlier, we then can get a representation of the full 1974 population.

The top panel of Table 2 shows the mean value for a number of variables in 1974 for the full population as well as for the sample both weighted and unweighted. In the bottom panel we present p-values for a comparison between the population means and the differently weighted samples. For all the variables, the weighted sample is representative of the full population. The unweighted sample, on the other hand, has significantly different means for every variable. This implies that as expected, the 1996 sample is not linked to a representative set of 1974 households, and instead certain types of households were more likely to be represented in the 1996 sample. The unweighted sample has a higher average family size,

---

[4] There were 3,825 households in 1974 that did not have any descendants in 1993. In this current paper we only focus on the 1974 households which have a descendant in 1993 because we cannot follow up those 3,825 households, although it is possible to examine and compare their characteristics with those of the households that do have descendants in order to determine whether their omission causes a bias.

which seems intuitive because a household with more family members is likely to have more descendants. On average, the households also have more cows and more rooms, both indicative of higher wealth. It seems that the 1996 sample is linked to a distinct set of 1974 households, that among other things are wealthier on average than the population, but using the weights calculated, we are able to make the sample representative of the 1974 population.

A researcher looking at the effect of the treatment on certain outcomes can use the linked sample of 1974 antecedent households in order to see what happens to the descendants of treated versus not treated households. By applying the weights, it is possible to obtain results that are representative for the full 1974 population. This is especially helpful when interested in heterogeneous effects of the treatment.

Another common method for calculating weights to correct for nonrandom sampling is to create a propensity score weight. This consists of using observable characteristics to calculate the probability that someone is picked to be in the sample.

We calculate propensity score weights for the 1974 population and compare them to the weights assigned using the sampling method. In order to do this, we take our sample of 1974 households linked to descendants in the 1996 sample and assign them a value of one for being linked, while all other 1974 households get a value of zero. This variable is the dependent variable in a logit regression. The controls used are observable characteristics of the 1974 households including highest education of anyone in the household, number of cows, number of boats, education of the head of household, the age of the head of household, the size of the family and the number of items owned by the family.[5] Table 3 shows the results of the regression used to calculate the propensity score. The coefficients from the regression are used to calculate predicted values for each 1974 household, which are equivalent to the probability that a certain 1974 household

---

[5] Information was collected on ownership of a lep, harrican, watch, radio and receipt of remittance. We also conducted the propensity score analysis with all of these variables as well as including the number of descendants. In a regular propensity score analysis this variable would not be available, but given that we have a full census in 1993, we have it and tried using it to see if the accuracy of the propensity score weights increased. There is no significant difference between the weights using the number of descendants variable and those not using it, so we only show the results for the propensity score weighting procedure where we do not include number of descendants.

is likely to be linked to the 1996 sample based on their observable characteristics. The weight is the inverse of this predicted probability.

The last column of Table 2 shows the results of the propensity score weighting, which can be compared to the sample weights, the scenario with no weights and the actual population means. Both our sampling weights and the propensity weights come very close to approximating the true population means. For most of the variables, the propensity weights are not significantly different from the population means, but they do differ significantly in the case of family size. Although the propensity weights are fairly representative of the full population, the difference in family size is worrisome because it could mean there are other unobservables that are also significantly different from the population averages. Therefore, our weights yield the most representative sample weighting structure.

A visual perspective on the effects of different weighting schemes is provided by Figure 3, which compares the cumulative distributions for the 1974 family size for the different weighting schemes. In addition to the basic measures presented in Table 2 (actual population, population weighted, unweighted and propensity score weighted) we add a measure of the estimate one would obtain if one used only the sampling weights from the actual 1996 sample to reweight the 1974 sample (sample weighted). This turns out to be a straightfoward calculation because of the fact that in the actual sample one household per bari is selected and sampling is independent across baris. In particular, based on the actual sample, the probability a particular antecedent household is picked is $(1 - \prod_j (1 - p_j))$ where j indexes all the 1996 households in the particular sample that are descendants of a particular 1974 household.[6]

The results are quite striking. As suggested by Table 2 the distribution for the actual population is very closely approximated by the distribution based on the weights simulated by redrawing the sample. The propensity scores does not do as well but is again quite close. On the other hand the sample weighted estimates coincide very closely with what one gets based on the unweighted data and both are quite far from the actual population. This result can be attributed to the facts that

---

[6] Note that if each 1974 household were linked to only one household in the sample, then the probability assigned in this way would just be the probability that the linked 1996 household is in the sample. These weights do not account for the fact that a 1974 household could theoretically enter the sample through any of its descendants, but only look at the particular descendants that it did enter the sample through.

one household per bari was picked in the 1996 sample and that many of the descendants of a particular 1974 household tend to live in the same bari. As a result the probabilities calculated from any particular sample gives very little sense of the likelihood of a particular 1974 household being picked across multiple draws of the sample. Specifically consider a household with five descendant households in one five household bari versus a household with 1 descendant in a five household bari. If that bari is sampled 100% of the time then in the former case the 1974 household will be picked 100% of the time but in the latter case it will only be picked 20% of the time. But since any given 1996 household is picked 20% of the time, one will assign the same probability of being picked if one bases, as is typically done, this probability on the probability that the actual sampled household was picked in a particular sample.

4 Correcting for Bias in Descendant Selection

The second problem with using the data without corrective weights is also related to the bari structure of the sample. Because baris were the unit that was randomly sampled, and only one household was picked from each bari, if two descendants from the same 1974 household were in the same bari, they would never both be picked to be in the sample. If, on the other hand, two descendants from the same 1974 household were in different baris, then it is possible that both could be picked for the sample, and even more so if they are in small or single household baris. We already mentioned how this could affect the representativeness of the 1974 population if the characteristics of 1974 households are correlated with the decision of their descendants to stay in the same bari or split off. In addition, if the decision to stay in the same bari as other descendants, split off into a new bari or join a different bari is correlated with the characteristics of the 1996 households, then it is not possible to accurately estimate average descendant outcomes, distorting the results of intergenerational analyses. This is especially a problem in evaluating the effect of the treatment because the treatment directly affects fertility and the number of descendants, which could lead to different patterns of household recombination that could affect the probability of being selected into the sample.

To illustrate the problem, suppose a 1974 household has three descendants and we are interested in the effect of the intervention on some outcome of descendants from this household. We need an accurate estimate of the average outcome for the descendants of the household, but in most cases all three descendants will not be in the sample. Now suppose that certain attributes determine whether households remain in the same bari or split off. For example, it is possible that the poorest descendant household of a family might choose to split off and look for better opportunities in a different location, while the two richer households remain in the same bari because they are already well off and would not want to leave their land, assets, network, etc. If this behavior were systematic in the population, it would mean that two high-outcome descendants, for example, would never both be in the sample, instead there would tend to be a richer and a poorer descendant. Thus, if the average outcome of descendants for particular households is calculated by taking the arithmetic average of the descendants that show up in the sample, then the sample will consistently underestimate the true average outcome of descendants. If the treatment led to an increase in the outcome of descendants, the treatment area would be impacted by this underestimate more than the non treatment area, leading to an underestimate of the program effect.

If household recombination is random, so that the probability of getting any combination of households with certain attributes is equal then there should be no bias, and the arithmetic average will be the average effect on descendants. This seems unlikely though, given that certain attributes such as wealth have been shown to play some role in household recombination.

To understand the problem more clearly, imagine the following scenario. Suppose there are two descendants from a 1974 household with 1996 outcomes c_1 and c_2. Theoretically we could see just c_1, just c_2 or both in the sample. In this example, household c_1 is never picked alone and there is a .5 chance of picking both c_1 and c_2 and a .5 chance of picking just c_2.[7] If we were to take the average of the households if they show up together and take the value of c_2 when we only have c_2, then we get the following expected outcome:

---

[7] This scenario might seem unlikely, but its purpose is to illustrate the more complicated case of several descendants where some live in the same bari and therefore the probability of picking two descendants living in the same bari is 0. Doing the example with more descendants makes the calculations messier and detracts from the point of simply illustrating why it is important to consider how descendants are weighted.

$$\begin{aligned} \mathbb{E}(c) &= 0 * c_1 + \frac{1}{2}\left(\frac{c_1 + c_2}{2}\right) + \frac{1}{2}c_2 \\ \mathbb{E}(c) &= \frac{1}{4}c_1 + \frac{3}{4}c_2 \neq \frac{1}{2}c_1 + \frac{1}{2}c_2 = \bar{c} \end{aligned} \tag{4}$$

In expectation, we are not getting the average for the descendants. If the probabilities were random with respect to the outcome of the households, then with a large enough sample, this inconsistency would average out. The problem arises if the probability is correlated with attributes of the 1996 households. For example, if the data consisted of two descendants for every antecedent and the probability of being picked is correlated with outcome so that c_2 is always the lower-outcome antecedent and c_1 is always the higher-outcome one, this would result in a lower estimate of the average outcome for descendants.

There are various ways to tackle this problem. In this case we could take c_2 if we only have c_2 and only take c_1 if we pick both c_1 and c_2, which would give us an expectation equal to the average. Yet in doing that, information on descendant c_2 would be thrown out if both descendants are in the sample. In addition, this is a solution for this particular set of probabilities. There are also more complicated probabilities in the data where there is a probability of seeing all three combinations.

Formally, we propose a method based on extension of the method used above to develop the 1974 measures. We then show that this method can be derived based on constrained minimization that is robust with respect to variation in the relationship between sampling probabilities and outcomes. We also explore the properties with respect to a simpler but less robust procedure. In particular, we wish to estimate

$$\Delta\bar{c}_{tN} = \frac{1}{N_{tx}} \sum_{i \in I_{tx}} \frac{1}{|A^{-1}(i)|} \sum_{j \in A^{-1}(i)} (c_{jt+1} - c_{it})^8 \tag{5}$$

As for Equation 2, we can construct an estimate of this quantity using only the sampled data and appropriate weights:

$$\Delta\hat{c}_{tN} = \frac{1}{N_{tx}} \sum_{i \in I_{tx}} \frac{1}{|A^{-1}(i)|} \sum_{j \in A^{-1}(i)} (c_{jt+1} - c_{it}) \frac{\mathbb{1}(j \in S_{t+1})}{\mathbb{E}(\mathbb{1}(j \in S_{t+1}))} \tag{6}$$

---

[8] $|A^{-1}(i)|$ denotes the size of the set of households descending from i.

It again follows that
$\operatorname*{plim}_{N\to\infty} \frac{\Delta\hat{c}_{tN}}{\hat{i}_{tN}} - \Delta\bar{c}_{tN} = 0.$
Note that (6) does not depend on the 1974 sampling probability. However, by dividing by this probability just after the first sum and multiplying by this probability after the second sum it can be seen that the appropriate estimates of average growth by initial characteristics is obtained by summing across descendant households in the sample that come from each 1974 household the scaled change in the outcome

$$(c_{jt+1} - c_{it}) \frac{\mathbb{E}(\mathbb{1}(i \in A(j)))}{|A^{-1}(i)|\, \mathbb{E}(\mathbb{1}(j \in S_{t+1}))} \tag{7}$$

and then combining the different 1974 households using weights derived from the resampled 1974 probabilities.

We now show how this estimate can be derived from a constrained minimization. The first criteria we want to meet is that any weights should lead to an outcome where the expected value for descendants' outcome is equal to the actual mean of the outcomes. The way this would look in the case of two descendants with outcomes c_1 and c_2 and conditional on observing at least one 1996 household with the probability p_a of picking just household 1 in the sample, probability p_b of picking both household 1 and household 2, and probability p_c of picking only household 2 in the sample is as follows:

$$\mathbb{E}(c) = p_a w_a c_1 + p_b(w_{b1} c_1 + w_{b2} c_2) + p_c w_c c_2 = \frac{1}{2} c_1 + \frac{1}{2} c_2 = \bar{c} \tag{8}$$

where w_a, w_{b1}, w_{b2}, and w_c are the weights applied to outcome 1 if it is only observed, outcomes 1 and 2 if both are observed, and outcome 2 if only 2 is observed, respectively.

There are many different possible ways of weighting the observations in order to get an expected value for descendants equal to the average. Some of these weights might lead to a large variance in the estimates, depending on the probabilities. Therefore, in addition to getting the correct mean income in expectation using our weights, we also want to minimize the variance from different probabilities of descendant combinations being picked for different antecedents. We want to minimize the following:

$$\begin{aligned}Z = [&var(p_a)(w_a^2c_1^2) + var(p_b)(w_{b1}c_1 + w_{b2}c_2)^2 \\ &+var(p_c)(w_c^2c_2^2) - 2cov(p_a, p_b)(w_ac_1)(w_{b1}c_1 + w_{b2}c_2) \\ &-2cov(p_a, p_c)(w_ac_1)(w_cc_2) - 2cov(p_b, p_c)(w_{b1}c_1 + w_{b2}c_2)(w_cc_2)]\end{aligned} \tag{9}$$

Both equation 8 and equation 9 depend on the actual mean and variance of the outcome values, but we do not have all of the outcome values. Therefore, the weights must work more generally and not be sensitive to the mean and variance of the outcomes. Again, this could be achieved in different ways. A sufficient condition for equation 8 to hold is that it holds for all outcomes, which can be ensured by taking derivatives with respect to the income values c_1 and c_2. In our simple two descendant example, this yields the following two equations, both of which need to hold in order for our first condition to be met:

$$\begin{aligned}p_aw_a + p_bw_{b1} &= \frac{1}{2} \\ p_bw_{b2} + p_cw_c &= \frac{1}{2}\end{aligned} \tag{10}$$

We apply a similar logic to our second condition and take second derivatives with respect to c_1 and c_2 in order to come up with the following objective function:

$$\min_{w_a, w_{b1}, w_{b2}, w_c} \frac{d^2Z}{dc_1^2} + \frac{d^2Z}{dc_2^2} \tag{11}$$

This ensures that the variation in the fraction of households in each sample has a small impact on the computed average income for descendants because the weights are applicable and minimize variance no matter what the actual incomes are. We calculate the weights by minimizing equation 11 conditional on equations 10. There are other, more complicated, criterion functions that could have been used, but we believe that this simple one still allows us to find weights that help to mitigate the potential bias arising from the bari structure. We will show how our weights using this procedure compare to not using weights and that indeed they can help to address the potential bias.

Solving the minimization problem, we find that the weights that minimize the variance of the estimates and in expectation yield the true average outcome are based on the probability of sampling a 1996 household. The combination in which a household appears (whether a descendant appears alone in the sample or if there are several other descendants in the sample from the same antecedent) does not affect the weight. This is surprising because as shown in our example in equation 4, the combination of descendants in a bari affects the expected value we get. Yet in trying to minimize the variance in a manner general enough to apply to all income values, the combination in which the descendants appear is no longer important. Nevertheless, the weight not only depends on the probability of being picked in 1996, but also on the total number of descendants. With the number of descendants in the denominator of the weight, those households who come from an antecedent with many descendants receive a smaller weight. Finally, the probability of the 1974 household *i* being represented in 1996 also factors in to account for the possibility that there is no descendant in the sample at all. Therefore, for our two descendant example the weights are:

$$
\begin{aligned}
w_a = w_{b1} &= \frac{\Pr(i)}{2 * \Pr(1)} = w_1 \\
w_c = w_{b2} &= \frac{\Pr(i)}{2 * \Pr(2)} = w_2
\end{aligned}
$$

where Pr(i) is the probability of the antecedent i of the household being represented by a descendant in the 1996 sample. We can generalize this result to assign a weight to every descendant j of a 1974 household *i* with M descendants:

$$
w_j = \frac{\Pr(i)}{M * \Pr(j)}
$$

This expression is exactly the formula derived above (Equation 6).

Without the outcomes of descendants to help assign the weights directly based on our original two specifications, it was necessary for us to come up with weights that are generalizable no matter what the incomes might be. Given that there are a number of ways we could have devised the weights, it is important to show that using the weights improves estimates. To do this, we have conducted a simulation to demonstrate how the weights compare to not using weights.

The simulation is a simplified case of our data in order to focus on the effect of the weights when the probability of being picked is correlated with the outcome, and how the performance of the weights depends on the extent of the correlation. We do not incorporate the sampling structure but instead look at how our weights perform in the case where we have 1000 antecedents and each has at least one descendant chosen for the sample. Therefore, here our Pr(i) is equal to 1 because each antecedent has probability 1 of having a descendant in the sample. [9]

In the simulated data, each antecedent has exactly two descendants and each of their descendants has randomly been assigned a log outcome from a normal distribution with mean 8.58 and variance 1.15.[10] This mean and variance were chosen as they were the mean and variance of one outcome in the data, log consumption for the 1996 sample.

We assigned probabilities for the following three events: household 1 is selected, household 2 is selected, both household 1 and household 2 are selected. The probabilities of being chosen are based on the random outcomes using a logistic function in order to ensure a correlation between outcomes and probabilities. We varied the size of that correlation by multiplying outcomes in the logistic expression times a coefficient δ, which is manipulated. The same δ is used for all

---

[9] Expanding this simulation to include the sampling structure that determines the probability that an antecedent household is selected does not change the results of how our weighting scheme compares to not using weights. This is because if we used the original sampling structure, we would then need to multiply times the 1974 weight of each antecedent in order to get the results for the population. But due to the fact that we had to create the 1974 weight based on the probability that an antecedent's descendant is in the 1996 sample, this is the same as Pr(i), and so multiplying times the inverse of this, which is our 1974 weight, we have:

$$\begin{aligned} W_{final}(j) &= w_{1974} * \frac{\Pr(i)}{M * \Pr(j)} \\ w_{1974} &= \frac{1}{\Pr(i)} \Longrightarrow \\ W_{final}(j) &= \frac{1}{\Pr(i)} * \frac{\Pr(i)}{M * \Pr(j)} = \frac{1}{M * \Pr(j)} \end{aligned} \tag{12}$$

Therefore, the final weight that is used is the same as our simplified weight in the simulation, so it is not necessary to complicate things by including the sampling structure.

[10] Although we only do the simulation with two descendants per antecedent household, the results are generalizable to more descendants and we have done some simulations including more than two descendants, but do not include the results here. We have also done simulations where we change the variance of the income variable, and this has also not affected the general result, so we omit those results.

three probabilities. The three probabilities are:

$$\begin{aligned} Pr(1) &= \frac{e^{\delta * c_1}}{e^{\delta * c_1} + e^{\delta * c_2} + e^{\delta * \bar{c}}} \\ Pr(2) &= \frac{e^{\delta * c_2}}{e^{\delta * c_1} + e^{\delta * c_2} + e^{\delta * \bar{c}}} \\ Pr(1\&2) &= \frac{e^{\delta * \bar{c}}}{e^{\delta * c_1} + e^{\delta * c_2} + e^{\delta * \bar{c}}} \end{aligned} \tag{13}$$

where $\bar{c}$ is the arithmetic mean of the two outcomes. The coefficient delta varies from 0 to 1 in .01 intervals. A coefficient of 0 implies that each event has a one third probability of occurring irrespective of outcome, so there is zero correlation between outcomes and probability. As the coefficient grows, the dependence between outcome and probability increases up to when the coefficient becomes 1, which gives the highest dependence between outcome and probability.

When the correlation is positive, if one household has a higher outcome than another, it will always have a higher probability of being selected alone, the probability of selecting both households will be next highest, and the probability of selecting the low-outcome household alone will be smallest. As the coefficient grows, this ordering of probabilities does not change, but the differences in probabilities become starker.

A sample of descendants is chosen based on the probabilities. One of the events is randomly chosen based on the probability of each event occurring.[11]Depending on the combination of descendants chosen for each antecedent, the weighted mean income is calculated based on the weights ($w_j = \frac{1}{2 * \Pr(j)}$). The mean outcome is also calculated with no weights, which entails taking the arithmetic mean if both descendants are in the sample, and taking the plain value of the descendant chosen if only one is in the sample. These two means are compared to the actual mean outcome for each descendant. Actual mean outcome is subtracted from the simulated mean outcome with and without weights and the absolute value is averaged to find the mean difference between actual and sample descendant outcome for the 1000 antecedents.

---

[11] This is done by assigning each event a piece of the unit interval equal to its probability, and then randomly choosing a number on the unit interval which determines the event based on which piece the number falls into.

In order to make sure the simulation is robust to outliers in the events picked, a set of events was chosen 500 times. The average of the absolute mean difference and squared error was calculated for each sample. This analysis was done with 1000 samples having different income values (and thus probabilities). This procedure was done for each δ from 0 to 1 in .01 intervals.

Figure 4 shows the sample average difference between the outcome calculated with our weights and the actual outcome, as well as the sample average difference between the outcome calculated without weights and the actual outcome.This is graphed for various deltas which represent how dependent the probabilities are on outcome. A delta of 0 signifies that the probability is not dependent on outcome at all, and a delta of 1 signifies a high degree of dependence between outcomes and the probabilities. The figure demonstrates higher variability in the average error when no weights are used versus when weights are used. Although the weights do not always perform better than not using weights, on average they are consistent in the level of error no matter what the correlation between outcome and probability, and this level is relatively low. The level of error ranges from almost none to almost .35 when no weights are used, while the error with the weights remains consistently under or close to .1.

If there is no correlation between outcome and the probability of certain combinations of households being picked, it would be better to use no weights. In that scenario, the mean descendant outcomes calculated using no weights are very close to the actual mean outcomes for the sample. When delta goes above .17, then the weights become better to use. This switch occurs for a relatively small delta, so if there is reason to believe a link exists between outcome and the probability of being chosen in the sample, then it is better to use the weights as compared to no weights.

The extent of the correlation can be tested using the census conducted in Matlab in 1996 that collected data from most of the households that were present in 1996. [12]. This survey has very limited data, but it does include several variables on household assets and household infrastructure which can be used to create an index as

---

[12] The MHSS sample was drawn on the 1993 census. However, we also use data from the 1996 census, because it contains various asset measures and household structure. It does not, however, included education

a proxy for household income. [13] The index is based on the importance of these assets and infrastructure for predicting income in the MHSS. Although the MHSS does not have an income variable, it provides data on consumption and the value of assets, which are regressed on the assets and infrastructure variables in 1996 that are analogous to the 1996 SES ones.[14] The coefficients from these regressions give two sets of weights for how important different assets are in predicting the income of a household. These weights are applied to the asset and infrastructure variables in the 1996 SES survey and create an index based on consumption and an index based on value of assets. These measures of consumption and wealth, used as proxies for income, can be used to calculate the correlation between income and the probability of being chosen for the sample. The probability values come from the simulations done for the first set of weights. The log probability of being picked to be in the MHSS is regressed on the log income index to get an elasticity. Using the consumption based index gives an elasticity of 0.339 significant at the .01 level, and using the assets value based index gives an elasticity of 0.402 significant at the .01 level as well.[15]

To put this in the context of where these coefficients fall in terms of δ, similar regressions were run using the simulated data. Regressions of the probability of being picked on the simulated income were run for each δ between 0 and 1 at .01 intervals. Figure 5 plots how the coefficient grows linearly as δ increases. The horizontal dashed and dotted lines plot the coefficient values based on the regressions using the consumption index and the asset index respectively. The vertical dash and dot line marks the δ at which the no weight and the weight lines crossed in Figure 4. The coefficient from the consumption regression crosses a little below where it becomes better to use the weights, while the coefficient from the asset regression crosses above where it becomes better to use the weights. This makes it difficult to make a definitive statement on whether it is better to use the weights.

---

[13] The variables used for the index are whether a household has a cow, boat, clock, or radio; whether it gets its drinking water from a tubewell; and whether the roof is made of tin.

[14] The assets used to compute the value of assets are homestead land, ornaments (gold, silver) savings in bank, television, radio, clock, electric fan, cycle, and furniture.

[15] These are the values from running a regression using all of the descendant households of the 1974 households that are present in 1996 and that have data in the SES in 1996. Given that there are 2,599 households in 1993 that are not present in 1996 (6 percent of households present in 1993), such data are not available on the full set of descendants for some 1974 households. The regressions were restricted to 1974 households for whom there was data on every one of their descendant households in 1996, and the results were very similar (.339 for consumption and .401 for assets value, both significant at the .01 level).

Nevertheless, due to the use of an index to proxy for consumption and an index to proxy for the value of assets, there is a possibility of measurement error. Testing this using the actual consumption and value of assets in the 1996 MHSS and the indexes created, there is evidence of classical measurement error. Assuming this is true of the full population data, it would imply our coefficients are biased downward.[16] The assets coefficient is already at a point that would imply that using the weights is better, and if the coefficients are biased downward, then it is likely the consumption coefficient would also be above the point marking indifference between using and not using weights. Therefore, in the case of the Matlab data, it makes sense to use the weights devised when conducting aggregate analyses.

Up to now, the discussion has focused on the performance of the weights on average for the sample. Except in the cases where there is very little correlation between income and the probability of being selected, the weights give a good approximation of the mean income of descendants on average for the sample. If we are interested in how they perform for individual households though, the average squared error is much bigger when using the weights as compared to not using weights. This is because as the sample gets bigger, the average income with the weights will be close to the expectation (in accordance with the law of large numbers), which due to the construction of the weights should be equal to the actual mean income. So with a sample of 1000, this holds true. But when looking at each individual antecedent and how the income calculated with the weights compares to the actual income, the squared difference is much bigger because the weights can cause some outliers. This is because chance means that sometimes even an event with a small probability will be picked, but that means it will have an extremely large and distorting weight.

Figure 6 shows how the weights compare to using no weights for individual antecedent households. The average squared error for an antecedent when not using weights is around 0.4, while the average squared error per antecedent starts out a little less than 9 and grows to almost 20 when using the weights. This is because, especially as the probabilities become more dependent on the income draws, there are more likely to be outliers with a very small or very large probability, which means a very large or small weight. Using such a large or small weight will give a more skewed average income than if no weight is applied. Nevertheless, even though for individual antecedents there is higher variability in the mean income calculated, looking at the whole sample, the very small incomes (due to very small

---

[16] Evidence of classical measurement error can be supplied by the authors upon request.

weights) will be balanced by the antecedents that get very large incomes due to large weights, and in this way the average income for the sample will approach the actual average income.

What this means is that it is important for a researcher to think about the type of analysis he or she is running in order to determine whether using the weights is appropriate. In the case where one is interested in average effects, such as running regressions, using the weights would lead to more accurate results (if there is a link between the probabilities of being picked and the variable of interest). If, instead, one is interested in the effects on certain quintiles of the population, which would involve using the average income to break people up into those quintiles, then the weights would distort the data. Therefore, it is important to be aware of the goal of any analysis and how using these weights might affect it in order to make sure that the weights are used correctly and are helping to improve the accuracy of the results rather than leading to additional bias or distortions[17].

5 Census Data

Before proceeding to the weighted data we take advantage of the population level data, inclusive of the 1996 census, to characterize overall economic and demographic mobility and to provide a benchmark that will serve to evaluate the effectivness of the different weighting schemes. The 1996 census data contains basic characteristics of households including the demographic structure and some household asset information but crucially, does not obtain information on education, thus necessitating the use of the sample weights to look at economic mobility.

We start by constructing an indicator that permits us to aggregate educational attainment across ages within households and to compare changes over time. Making use of the education data we determine the mean and standard deviation of completed schooling for children aged 6-16 in the MHSS using the household cross-sectional weights. We then substract from each individuals education, both in the 1974 and 1996 data, the mean schooling and divide by the standard deviation of schooling from 1996 calculated by age to create an educational z-score. We

---

[17] Although in the empirical section we break up the 1974 data into thirds based on income, we use the full population and not the weights to do this. We only use the weights in running the regressions

then aggregate at the household level and year to obtain, in effect, a household level z-score of child education that reflects the relative intensity of child education of each household. In addition to the education data we have the consumption index constructed using the sample data that was described in the previous section and measures of the family size in 1974 and 1996 census. For purpose of the regression analysis we divide 1974 household z-scores, the 1974 family size, and the income index into three categories in order to condition on possible pretreatment differences in characteristics to look at possible heterogeneity in different outcomes by treatment status. We use the middle category in each case as the reference category.

Plots of the cumulative distribution of the educational index by treatment and comparison and year are presented in Figure 7. The 1974 data are based on the census and the 1996 data are based on cross-sectional sampling weights from the MHSS. Not surpisingly educational intensity in 1974 was about one standard deviation (relative to the 1996 base) less than it was in 1996. We also see differentials in education between treatment and comparison both before and after the program, though the differences appear to be somewhat larger in 1996. This result is important because it suggests that there were preexisting differences between treatment and comparison area in education in 1974 that should lead one to question the validity of inference about the effects of the treatment program on educational differences in 1996 without controlling for preexisting differences. As noted above, this is one important reason for looking at household education in 1996 conditional on the education of 1974 antecedent households. [18]

Figure 8 provides a lowess plot of the average antecedent-descendant change in family size by household education z-score and treatment status. Overall family sizes dropped by 1-2 persons per household between 1974 and 1996 and this drop was lower among the lower educational intensity households in 1974 than among the higher educational intensity households. There was also a uniform drop by educational intensity when comparing treatment and comparison households. Consistent with the notion that the family planning program lead to rapid increase in contraceptive use and decline in fertility, the treatment households

---

[18] These differences were noted in Roy and Foster 1996 and are attributable to the presence of the Matlab regional capital being in the southern part of the treatment area. That paper shows there were not preexisting differences onces this area was excluded. If one were not directly interested in household educational mobility then it may be possible to address the issue by conditioning, for example, on educational differences at the level of the village

declined by about half a person more per houseohlds than did the comparison households. The uniformity is striking, evidently the intervention was quite ubiquitous in its effect not being differentially adopted by households with higher education. Indeed a regression analysis reported in Table 4 confirms this impression from a statistical perspective. On average the treatment program lead to a significant .57 decline in family size and none of the interaction terms between treatment and baseline status are different from zero individually or collectively. We also see from the baseline characteristics evidence of convergence in family size–large families lost on average four members and small families increased in size by about 2. Family sizes decline the most among higher education households but there are no significant differences in family size by baseine consumption per capita.[19]

Of course, family sizes may have also been affected by changes in household division. If, for example, the treatment program had lead to lower levels of household division because, say of lower crowding, then one would expect the results in Figure 8 to underestimate the effects of fertility on family size. This seems not to be the case, however. Figure 9 presents the number of descendant households per antecedent household by educational intensity and treatment status. There does seem to be a relationship between educational intensity and household division, with the highest number of descendant households among households with an educational intensity roughly one standard deviation below the mean in 1996 showing the highest average number of descendants–close to 4. However, there is only a small difference in the number of descendant households between treatment and comparision areea. Again Table 4 provides statistical evidence on this point. Net of other factors, higher education households had fewer descendant households, larger families had more descendant households, and wealthier households had fewer. But the overall treatment effect is not statistically significant (column 5) and the only interaction effect that is statistically significant is for lower consumption households, which appear to have more descendant households in the treatment area.

Figure 10 provides a similar perspective with respect to change in consumption per capita as measured by the respective indices. Overall there seems only a small difference in consumption growth between treatment and comparison households

---

[19] Although Figure 8 suggests a U shaped relationship between household size and education, there are very few households with a Z-score between 1 and 2, therefore the highest education group in 1974 encompasses the households with the highest decrease in education that have a Z-score between .5 and 1.

given educational intensity in 1974. This result suggests that the evident reductions in family size by 1996 have not lead to significant increases in the per capita availability of certain assets. Of course, this result should be interpreted with some caution because the consumption index for 1974 is based on the observed relationship between consumption per capita and assets and family size in 1996. Suppose the true coefficient on the coefficient on family size fell from 1974 to 1996 . Then the estimated difference between 1996 and 1974 will be negatively correlated with family size in 1974. If family size is correlated with educational investment this could affect the observed relationship between estimated consumption growth and educational investment by treatment status. One partial fix would then be to condition on 1974 family size as well as 1974 educational investment. The regression in Table 4 which controls for each of the initial variables does indeed give a different picture. Once one conditions on initial consumption and initial family size as well as educational investment one sees small (1.7%) differential consumption growth in the treatment area households. Again, this difference seems, like family size, to be fairly ubiquitous with none of the individual interaction coefficients in column (5) being significant at greater than the 10% level. In terms of the level effects we see greater consumption growth in the higher and lower education households, in the larger families in 1974, and among those households with lower baseline consumption.

6 Household Education Using the Sample Weights

Before proceeding to the analysis of educational outcomes, which requires the use of sampling weights, we consider the application of our different weighting schemes to family size change which has the advantage that we can compare different weighting schemes to those at the population level. We do this using the approach from Table 4 but exclude the interaction terms as the family size changes seemed to be similar across the different strata. The results appear in Table 5. The first column replicates column (2) of Table 4. The second column implements our preferred scheme. A combination of 1974 and 1996 probabilities and the total number of descendant households are use to average results across descendant households and then these are aggregated using the 1974 resampled weights. The third column is analogous to our propensity score approach and uses the predicted number of descendant households based on a regession that includes only the 1974 asset variables. The basic idea is to find an approach that would be effective in the presence of sampled data but without full information on the actual number of descendant households for each antecedent. The fourth column constructs the simple within sample average of education of the descendants but then weights using the 1974 sample weights. The fifth column constructs a weighted mean among the sampled descendants using the 1996 cross-sectional weights and then weights antecedents using the 1974 sample weights. The sixth column is the same as the fifth except that it does not adjust using the resampled weights. Finally the seventh column works with the unweighted data. Estmates are clustered using the 1996 household to reflect the fact that each 1996 household, once sampled, can contribute multiple antecedents and thus contributes multiple observations.

As is evident from the table, the formal weight procedure matches quite closely the results from the full population with a treatment coefficient of -.567 on treatment versus -.589 in the population. The standard error is about three times larger reflecting the much smaller size of the sample. The predicted sample is in the same ballpark but smaller and less significant. Evidently there is significant loss of information when using the predicted number of descendants rather than the actual number. The other measures provide a somewhat mixed bag. The combination of 1974 and 1996 weights gives the closest estimate to -.589 among these remaining measures. The only difference with our preferred estimate is that it uses averages over sampled descendants rather than the actual number of descendants. The similarity is perhaps not surpising given the simulation results above. Recall that the correlation between the consumption index and the weighting probabilities in the data is right at the crossing point in 4. Two of the estimated treatment effects are too large in absolute value: the simple 1974 weights and no weights. Both of these approaches ignore the differences in 1996 sampling probabilities that are attributable to bari size. Thus the gap seems to be attributable to differences in bari size that may reflect differences in household division that were higher for high consumption households in the tratment area as illustated in Table 4. On the other hand, the 1996 weights alone seem to result in the an underestimate of the treatment effect.

We now turn to the analysis of educational mobility. Figure 11 presents the 1996 household average child educational z-score by 1974 education and treatment status and is constructed using the formal procedure defined above given the absence of educational attainment data in the 1996 census. The hashed lines indicate confidence intervals. We see, consistent with the results reported previously, that there was an overall increase in child education in the treatment area relative to the comparison area for households with the same 1974 educational attainment. However, this effect appears to dissipate at the lowest levels of education. Evidently households with low levels of education in 1974 did not differentially benefit from the treatment program even though there was a comparable decline in family size for these households as shown above.

The results are confirmed in the statistical analysis that appears in Table 6. These specifications include overall treatment status and the interaction between treatment status and low initial education.[20] Other interaction terms were not significant and have been omitted so that we focus on comparisons across the different estimation methods. The point estimates suggest that except among low education households in 1974 there was a .216 standard deviation increase in the education among the 1996 households. There is by contrast no effect (and the point estimate is negative) in the lowest education households. Thus the treatment program was in a sense regressive because the higher education households benefited differentially from the program. In terms of level effects we see a .364 standard deviation increase in education among the higher educational households in 1974 and a very strong gradient in terms of consumption with high consumption households gaining .204 standard deviations and low consumption households losing -.167 relative to the mean education. These data suggest substantial increases overall in education and relatively little cross-group mobility in education on average.

Overall the different weighting schemes seem to give somewhat different conclusions. The method using predicted descendants is quite comparable to that using total descendants. The other weighted methods tend to yield a somewhat smaller estimate of the treatment effect with the two estimates using the 1996 weights alone or with the 1974 weights being substantially smaller. Interestingly, the no weight estimates do rather well. While this gives one some pause about the nature of this exercise it is worth noting that there is no way to predict that in practice. We expect it is an issue of the fact that the 1974 weights tend to bias upward the estimates of the treatment effect and the 1996 weights to bias it downward. The combination seems to come in about right, at least given the particular correlations between weights and the educational outcome variable being examined here.

---

[20] Note that the number of observations is lower because only those households with children ages 6-16 for whom we calculate education are included.

7 Conclusion

We started out wanting to answer the question of what are the long term effects of a maternal and child health and family planning program on economic mobility. In order to properly conduct this analysis, though, we had to think through the broad question of how to devise the appropriate weights for panel data when there might be bias in the selection of the sample, and we created those weights based on the MHSS and HDSS data from Matlab, Bangladesh. We first laid out the various issues that arise with the Matlab data due to the post-1978 selection of the MHSS sample. Although this is a problem specific to the Matlab data, it is one that might apply in any of the other HDSS sites where an intervention was conducted on a sample of the population, or only a sample of the HDSS population was later tracked after an intervention in the region. It could arise even in the case of regular panel data if the formation and recombination of households combined with the choice of descendants picked to be surveyed leads to selection bias in the sample that is followed up. Therefore, in the case of any development intervention where there are such data limitations, we have created a possible framework for weights that can help to mitigate the bias.

We devised a procedure to help solve the two main problems with the MHSS/HDSS data. For the first, to make the 1996 sample representative of the 1974 population, we used the nature of the HDSS data which allowed us to mimic the process that had been used to create the original sample in order to come up with probability weights. Even in the case of panel data where the full population is not available to conduct this sort of resampling procedure, propensity score weights also give extremely good results in helping to correct the sampling bias. For the second problem, we have found a formula for weights that can be universally applied in the case of multiple descendants where not all descendants have the same probability of being picked. Nevertheless, the application of these weights is not advisable if there is no correlation between the probability of being selected and the characteristics of interest, or if the analysis is not focused on aggregate data.

Using the sample weights to look at the main question of interest, we found that the maternal and child health and family planning program led to a differential increase in education among relatively high education households but a farily comparable decline in family size across the different strata. We can only speculate about the differences in effects. The most plausible explanation may be that households with lower taste for education (and or higher demand for child labor) reduced their family size but given demand for household child labor this increased the per child demand for household labor. Another possibility is that somehow the treatment effects differentially affected the process of household recombination. One reason that we might be skeptical of this conclusion is the relatively small differences across educational strata in treatment effects for family size, consumption and the number of descendant households.

As discussed, there are many similar interventions where baseline surveys and subsequent household composition changes of the study population are not available. Employing similar weights to the ones we have created here, which take into account the various issues created by the process of household formation and recombination, could lead to significant improvements in the evaluation of these development interventions.

References

Arunachalam, Raj and Suresh Naidu (2008). "The Price of Fertility: Marriage Markets and Family Planning in Bangladesh". In: *University of California, Berkeley*.

Asadullah, M Niaz (2012). "Intergenerational wealth mobility in rural Bangladesh". In: *Journal of Development Studies* 48.9, pp. 1193–1208.

Barham, Tania (2012). "Enhancing cognitive functioning: Medium-term effects of a health and family planning program in Matlab". In: *American Economic Journal: Applied Economics* 4.1, pp. 245–273.

Becker, Gary S and H Gregg Lewis (1974). "Interaction between quantity and quality of children". In: *Economics of the family: Marriage, children, and human capital*. UMI, pp. 81–90.

Fauveau, Vincent (1994). "Women, Children and Health". In: *ICDDR, B special publication* 35.

Field, Erica and Attila Ambrus (2008). "Early marriage, age of menarche, and female schooling attainment in Bangladesh". In: *Journal of Political Economy* 116.5, pp. 881–930.

Fitzgerald, John, Peter Gottschalk, and Robert A Moffitt (1998). *An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics*.

Foster, Andrew D (1993). "Household partition in rural Bangladesh". In: *Population Studies* 47.1, pp. 97–114.

Foster, Andrew D and Mark R Rosenzweig (2002). "Household division and rural economic growth". In: *Review of Economic Studies*, pp. 839–869.

Hausman, Jerry A. and David A. Wise (1981). "Stratification on Endogenous Variables and Estimation:The Gary Income Maintenance Experiment". In: *Structural Analysis of Discrete Data with Econometric Applications*. Ed. by C.F. Manski and D.L. McFadden. MIT Press.

Joshi, Shareen and T Paul Schultz (2007). *Family planning as an investment in development: evaluation of a program's consequences in Matlab, Bangladesh*. Tech. rep. IZA Discussion Papers.

— (2013). "Family Planning and Women's and Children's Health: Long-Term Consequences of an Outreach Program in Matlab, Bangladesh". In: *Demography* 50.1, pp. 149–180.

Maitra, Pushkar (2003). "Schooling and educational attainment: evidence from Bangladesh". In: *Education Economics* 11.2, pp. 129–153.

Moffit, Robert, John Fitzgerald, and Peter Gottschalk (1999). "Sample attrition in panel data: The role of selection on observables". In: *Annales d'Économie et de Statistique*, pp. 129–152.

Phillips, James F et al. (1982). “The Demographic Impact of the Family Planning–Health Services Project in Matlab, Bangladesh”. In: *Studies in Family Planning*, pp. 131–140.

Rahman, Omar et al. (1999). “The 1996 Matlab Health and Socioeconomic Survey: Overview and User’s Guide”. In: *RAND Corporation Draft Series DRU-2018/1-NIA*.

Rosenzweig, Mark R and Kenneth I Wolpin (1980). “Testing the quantity-quality fertility model: The use of twins as a natural experiment”. In: *Econometrica: journal of the Econometric Society*, pp. 227–240.

Roy, Nikhil and Andrew D Foster (1996). *The Dynamics of Education and Fertility: Evidence from a Family Planning Experiment*. Tech. rep. University of Pennsylvania.

Schultz, T Paul (2009). “How Does Family Planning Promote Development? Evidence from a Social Experiment in Matlab, Bangladesh 1977-1996”. In: *Yale University, Economic Growth Center, New Haven, Conn*.

Figure 1: Sampling Probabilities with Household Recombination

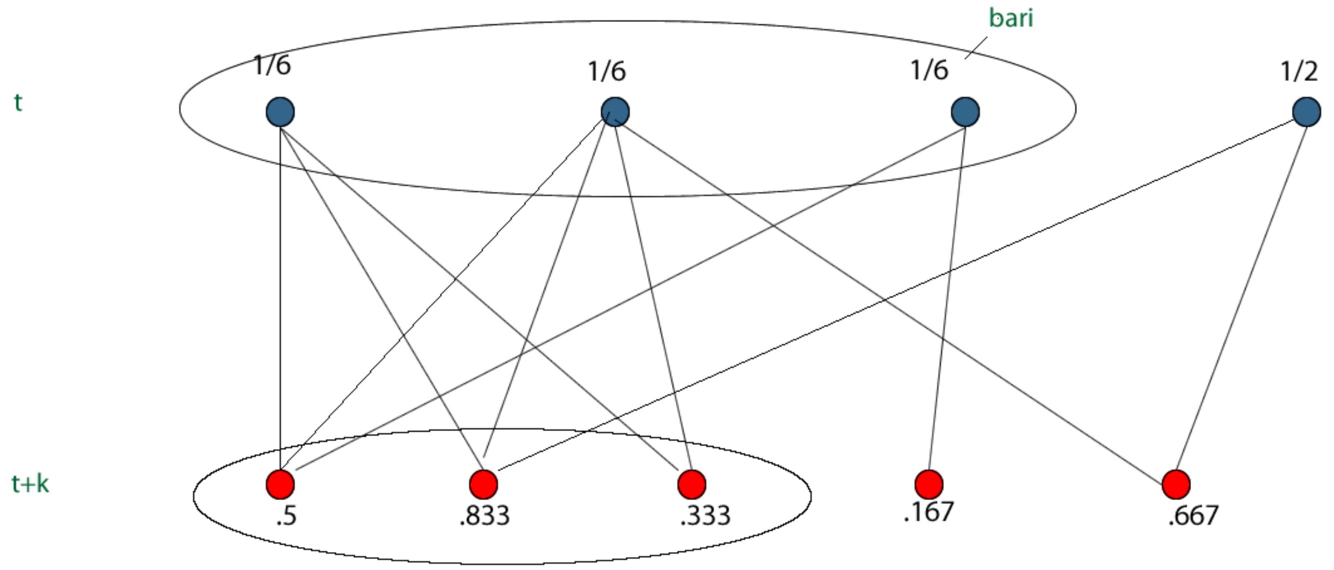


Table 1: Number of Links Between Selected 1974 and 1996 Households

	1996 Households										
	A11*	B11	C11*	D11	D12	D13	E11	C12	F11	B12	B13
1974 Households											
A01	3	4	0	0	0	0	0	0	0	0	0
D01	3	4	7	5	4	4	7	4	5	1	4
G02	3	4	0	0	0	0	0	0	0	0	0
D02	0	0	7	5	4	0	0	4	5	3	4
D03	0	0	1	3	3	0	0	0	5	0	0
C01	0	0	7	0	0	0	0	4	0	0	0
C02	0	0	7	0	0	0	0	4	0	0	0
E01	0	0	0	0	0	1	7	0	0	0	0
H01	0	0	0	0	0	0	0	0	0	1	2
J01	0	0	0	0	0	0	0	0	0	3	0
K01	0	0	0	0	0	0	0	0	0	3	4
G01	0	0	0	0	0	0	0	0	0	3	4
L01	0	0	0	0	0	0	0	0	0	3	0
B01	0	0	0	0	0	0	0	0	0	1	2

*Household was in the 1996 MHSS

Figure 2: Distribution of Fraction of Antecedent Households
in Treatment Area Among MHSS1 Households

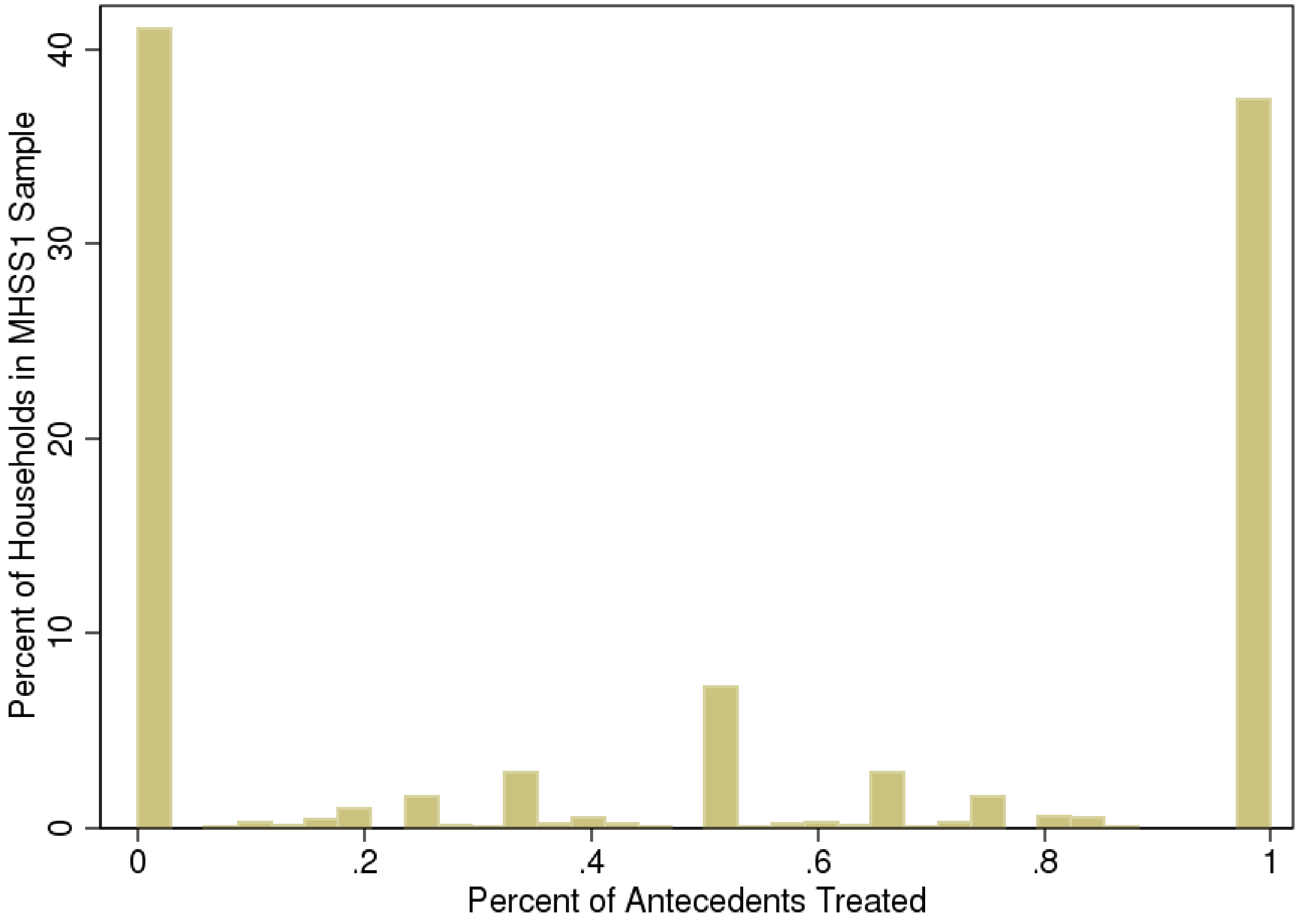


Table 2: Different 1974 Sample Weights Compared to the Full 1974 Population

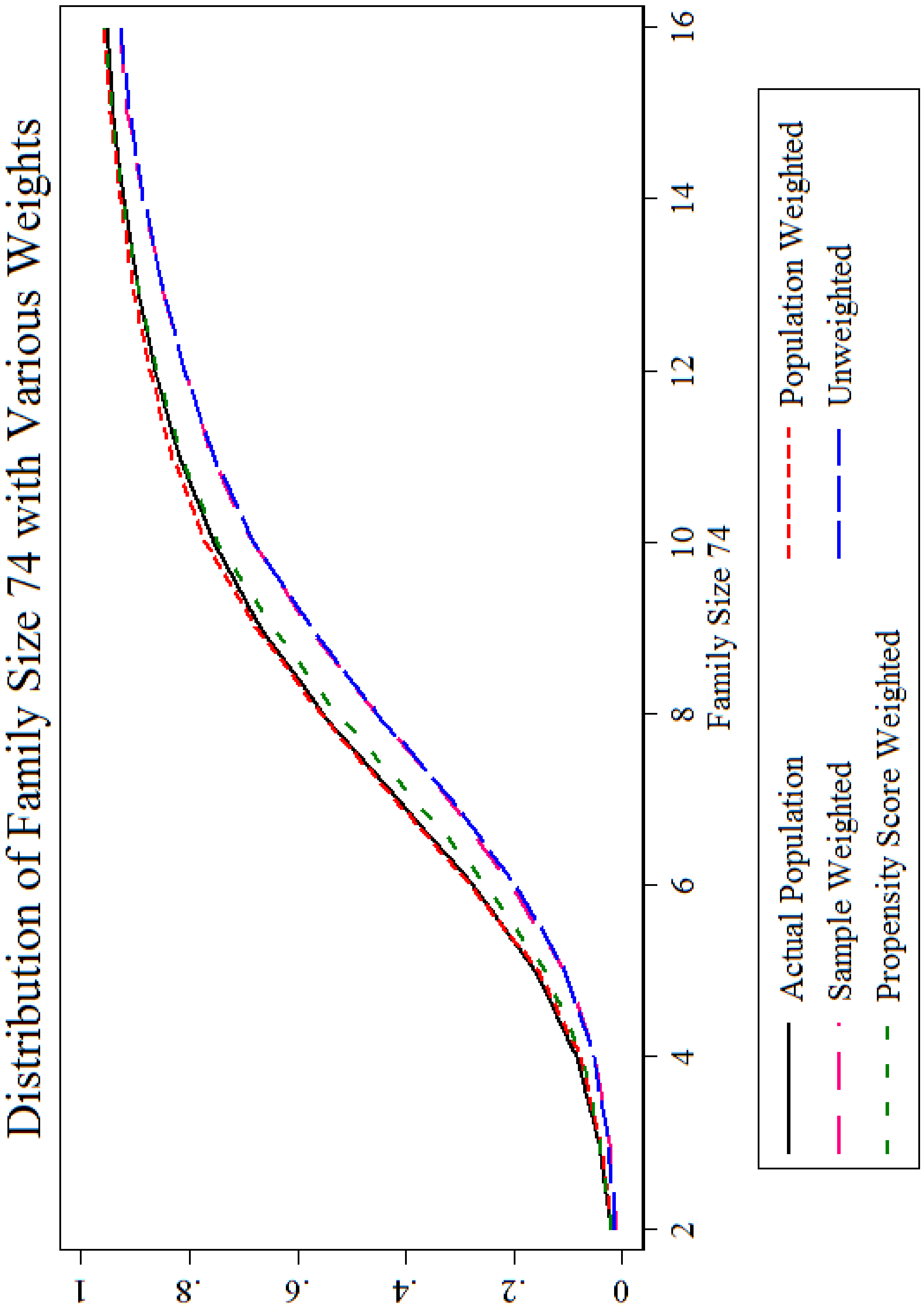
	(1)	(2)	(3)	(4)
	Mean Values for 1974 Population and Weighted Samples			
	Full 1974 Population	Our Resampling Weights	No Weights	Propensity Score Weights
Highest Edu	4.136	4.234	4.259	4.182
Number of Cows	1.158	1.180	1.353	1.174
Edu of Head	2.272	2.276	2.137	2.288
Age of Head	45.75	45.72	46.74	45.85
Family Size	6.095	6.127	6.811	6.185
Num of Rooms	1.219	1.213	1.284	1.215
Observations	24,788	5,319	5,319	5,309
Weights	24,788	24,029	5,319	24,594
	P Values for Difference between Full Population and Sample			
		Our Resampling Weights	No Weights	Propensity Score Weights
Highest Edu		0.260	0.030	0.427
Number of Cows		0.504	0.000	0.498
Edu of Head		0.953	0.003	0.752
Age of Head		0.939	0.000	0.611
Family Size		0.544	0.000	0.015
Num of Rooms		0.544	0.000	0.580

Table 3: Propensity Score Regression

VARIABLES	Dep Var: Dummy=1 if Linked to MHSS1
Highest Edu in Household	-0.0110*
	(0.00658)
Family Size	0.130***
	(0.00693)
Articles Owned	-0.00129
	(0.00181)
Number of Cows	0.0126
	(0.0111)
Number of Boats	0.0159
	(0.0290)
Edu of Head of Household	-0.0239***
	(0.00708)
Age of Head of Household	-0.000923
	(0.00127)
Constant	-1.991***
	(0.0620)
Observations	24,749

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 3



Distribution of Family Size 74 with Various Weights

Figure 4: Average absolute difference between sample and actual descendant income means for different levels of correlation

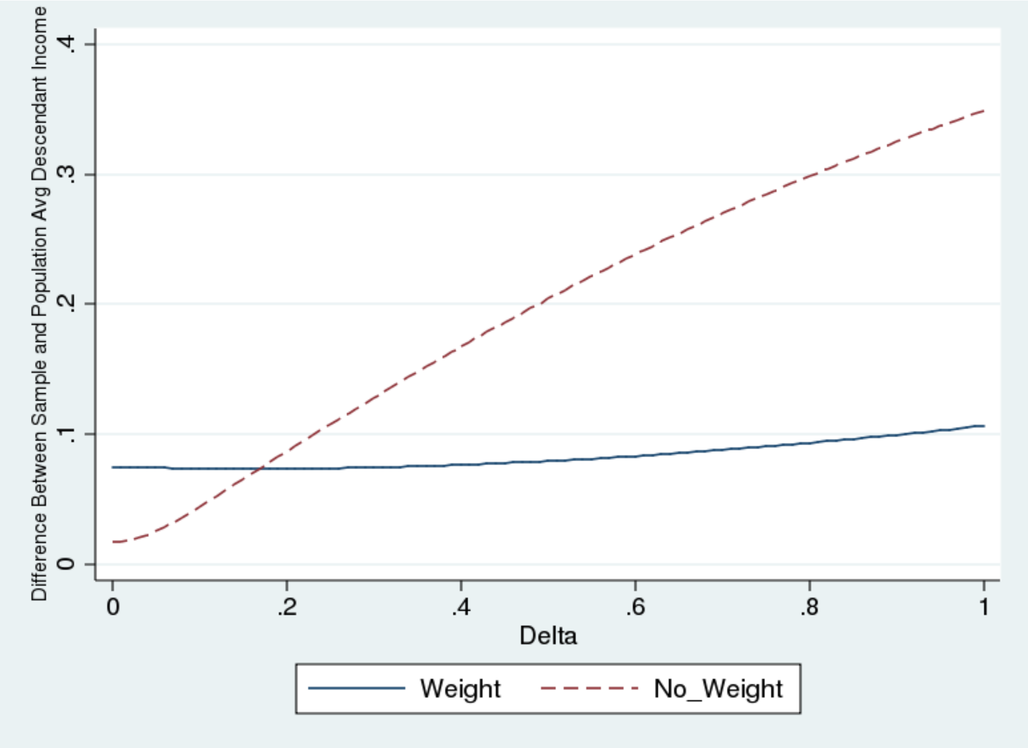


Figure 5: Coefficients of Probability Regressed on Income for the Simulated Data and the 1996 Census Data

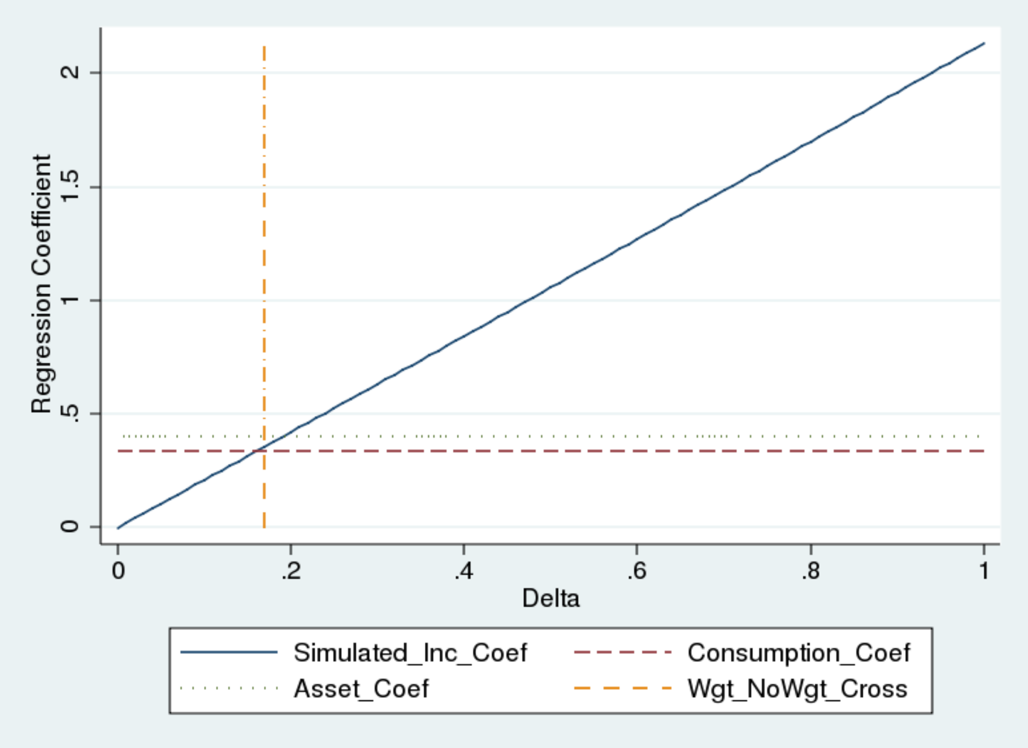


Figure 6: Average squared difference between sample and actual descendant incomes

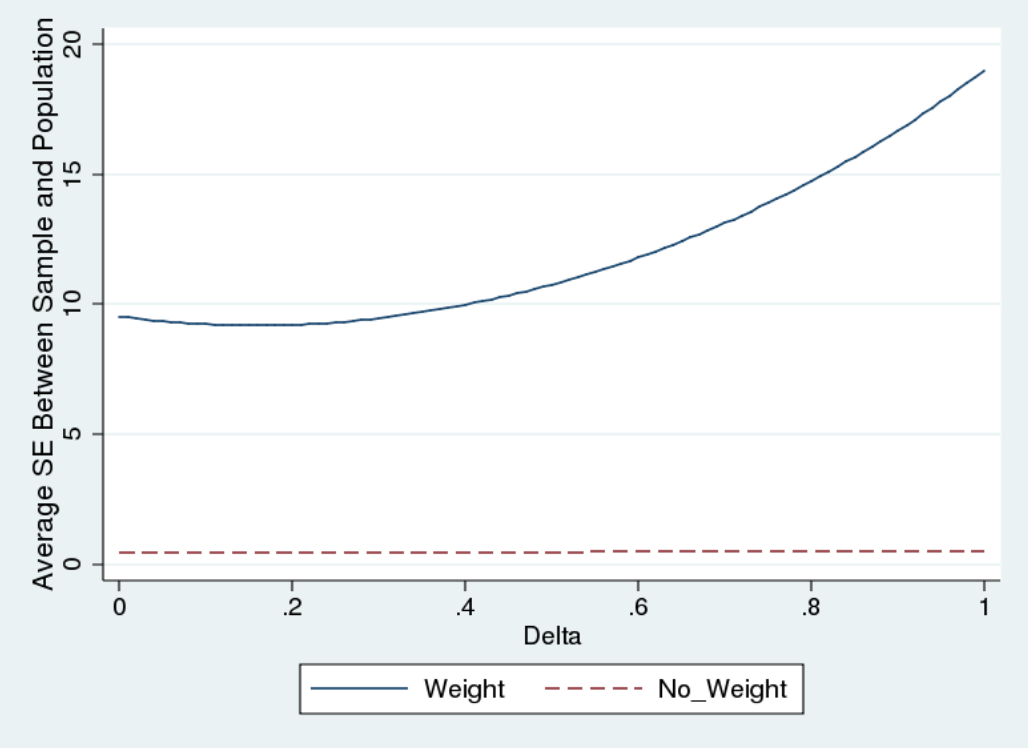


Figure 7

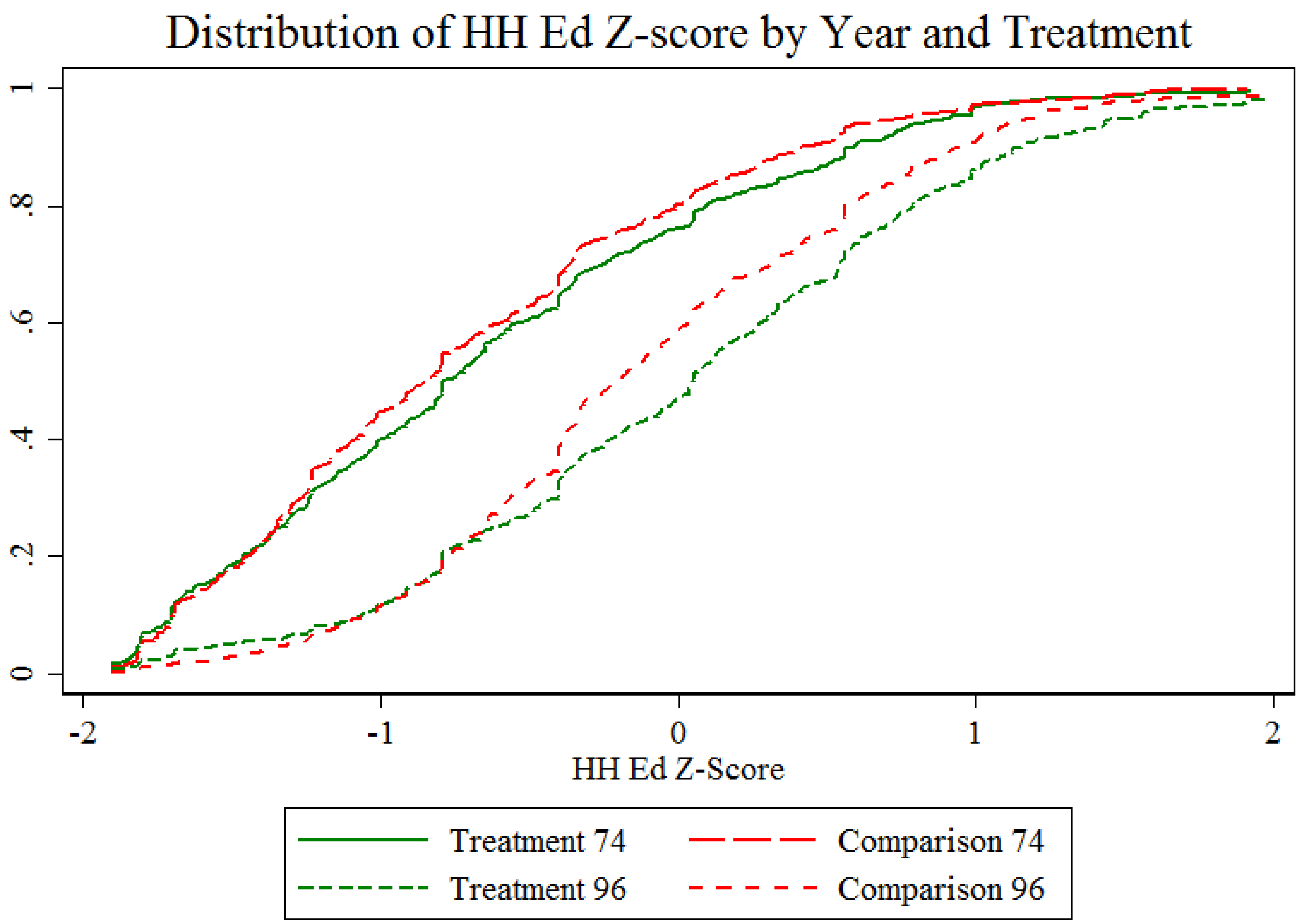

Distribution of HH Ed Z-score by Year and Treatment
0
.2
.4
.6
.8
1
-2
-1
0
1
2
HH Ed Z-Score
Treatment 74
Comparison 74
Treatment 96
Comparison 96

Figure 8

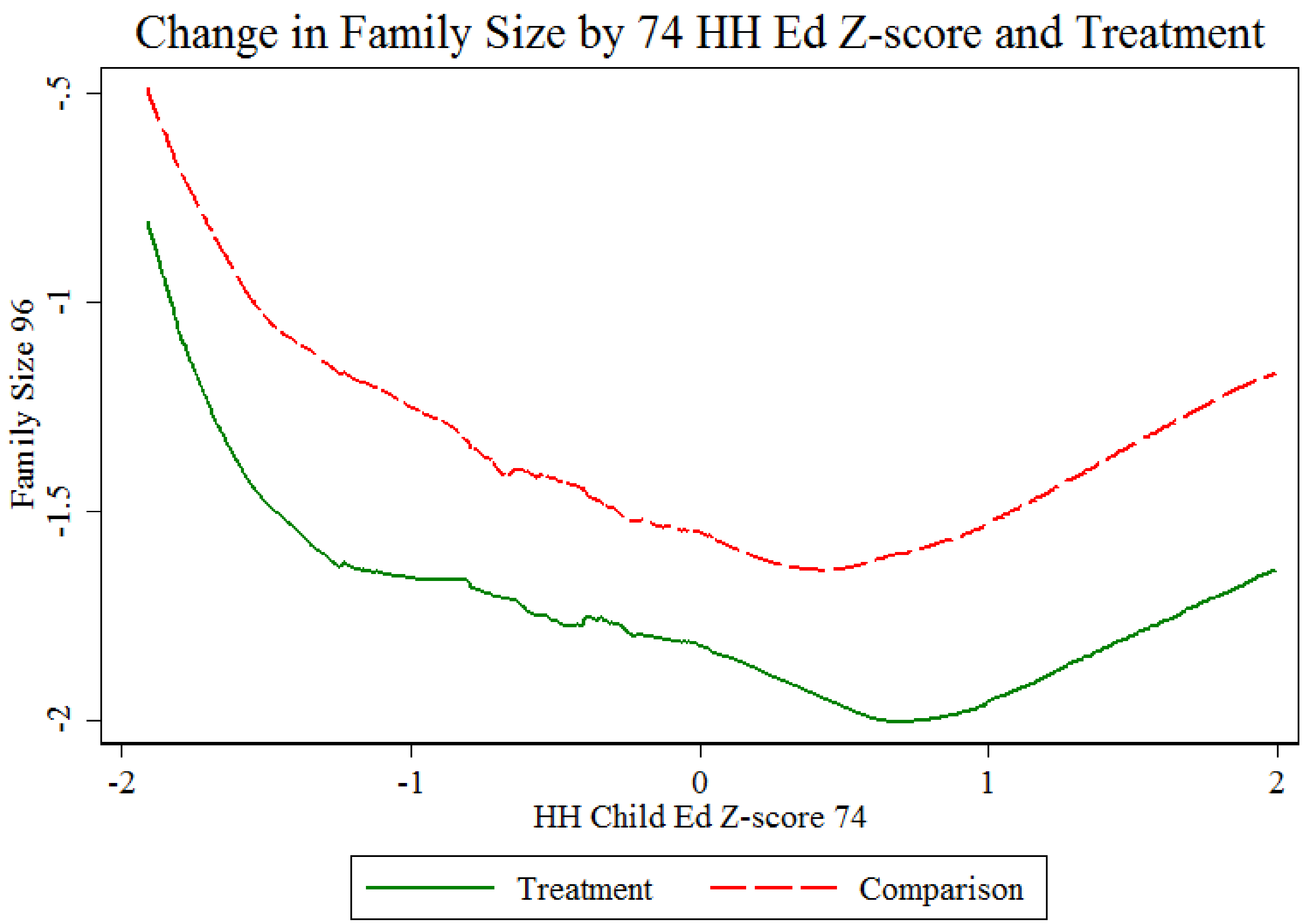

Change in Family Size by 74 HH Ed Z-score and Treatment
Family Size 96
-.5
-1
-1.5
-2
-2
-1
0
1
2
HH Child Ed Z-score 74
Treatment
Comparison

Figure 9

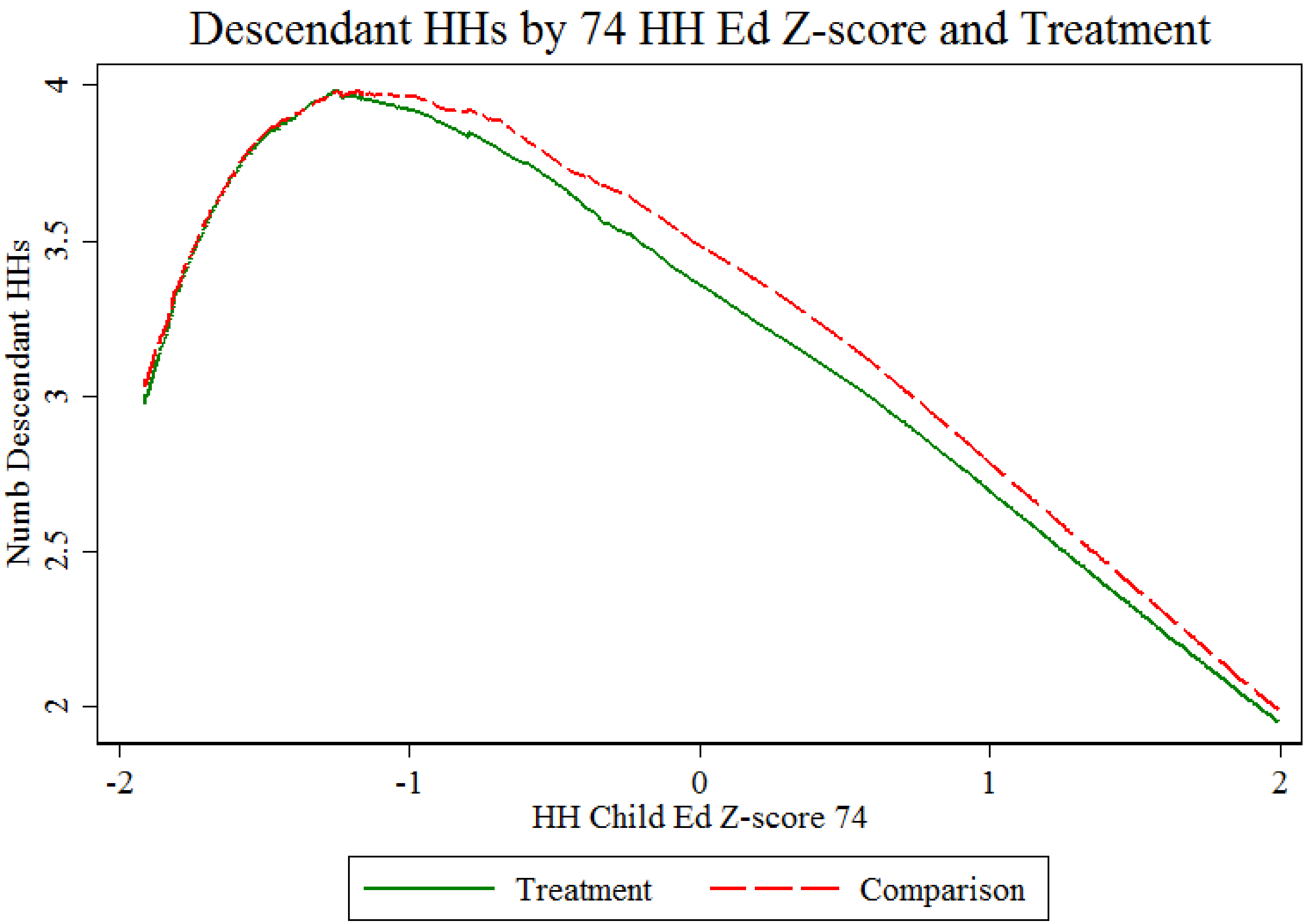

Descendant HHs by 74 HH Ed Z-score and Treatment
Numb Descendant HHs
4
3.5
3
2.5
2
-2
-1
0
1
2
HH Child Ed Z-score 74
Treatment
Comparison

Figure 10

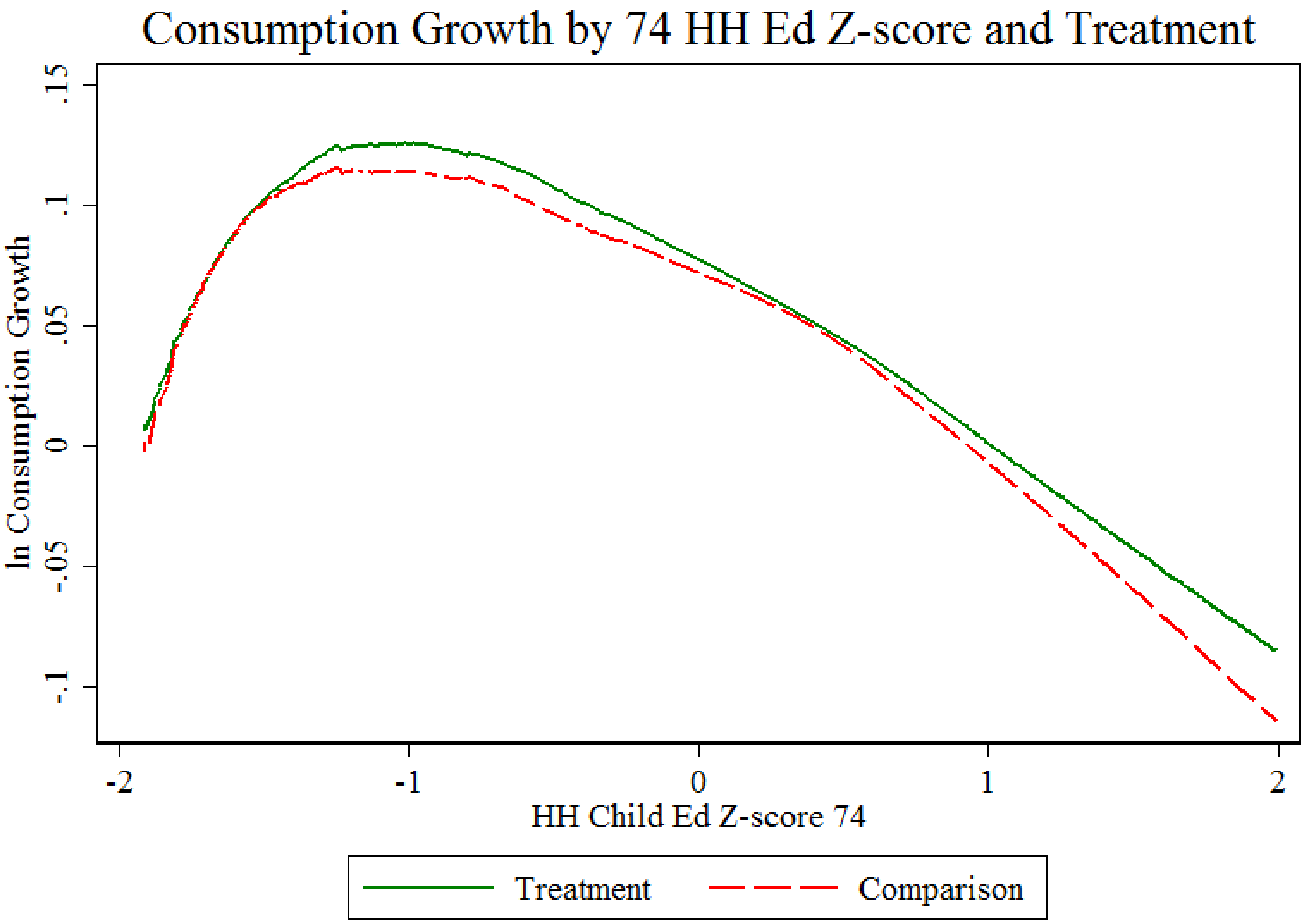

Consumption Growth by 74 HH Ed Z-score and Treatment

Table 4

Population Estimates by HH Child Z-score and Treatment

VARIABLES	(1) Change in Family Size	(2) Change in Family Size	(3) Descdant HHs	(4) Descdant HHs	(5) Consumption Growth	(6) Consumption Growth
Ed Low	-0.0461	-0.122*	-0.0223	0.0326	-0.0343***	-0.0416***
	(0.108)	(0.0729)	(0.0716)	(0.0493)	(0.00563)	(0.00392)
Ed High	-0.391***	-0.388***	-0.495***	-0.517***	0.0417***	0.0370***
	(0.140)	(0.0925)	(0.0775)	(0.0509)	(0.00649)	(0.00447)
F Size Low	1.755***	1.899***	-0.588***	-0.559***	-0.00152	-0.00509
	(0.129)	(0.0858)	(0.0799)	(0.0530)	(0.00827)	(0.00572)
F Size High	-4.352***	-4.343***	1.299***	1.221***	0.0195***	0.0241***
	(0.103)	(0.0689)	(0.0621)	(0.0434)	(0.00525)	(0.00365)
Cons Low	-0.702***	-0.730***	0.379***	0.538***	0.178***	0.180***
	(0.118)	(0.0815)	(0.0739)	(0.0517)	(0.00571)	(0.00397)
Cons High	0.0340	0.0677	-0.368***	-0.326***	-0.228***	-0.236***
	(0.121)	(0.0795)	(0.0736)	(0.0476)	(0.00651)	(0.00445)
T x Ed Low	-0.151		0.115		-0.0145*	
	(0.145)		(0.0972)		(0.00778)	
T x Ed High	-6.84e-05		-0.0339		-0.00878	
	(0.183)		(0.101)		(0.00881)	
T x F Size Low	0.273		0.0550		-0.00720	
	(0.171)		(0.105)		(0.0115)	
T x F Size High	0.0153		-0.147*		0.00896	
	(0.137)		(0.0853)		(0.00725)	
T x Cons Low	-0.0618		0.312***		0.00368	
	(0.160)		(0.106)		(0.00786)	
T x Cons High	0.0597		0.0808		-0.0147*	
	(0.160)		(0.0976)		(0.00885)	
Treat	-0.570***	-0.589***	-0.0931	0.0122	0.0229***	0.0169***
	(0.146)	(0.0727)	(0.0941)	(0.0459)	(0.00796)	(0.00373)
Constant	1.284***	1.293***	3.162***	3.108***	0.0691***	0.0723***
	(0.114)	(0.0866)	(0.0715)	(0.0554)	(0.00590)	(0.00462)
Observations	19,319	19,319	19,319	19,319	19,319	19,319
R-squared	0.249	0.249	0.093	0.093	0.336	0.336

Robust standard errors in parentheses

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

Table 5

Family Size Change by 74 Conditions and Treatment

VARIABLES	(1) Population	(2) Formal	(3) Predicted	(4) 74 Weights	(5) 74/96 Weights	(6) 96 Weights	(7) No Weights
Ed Low	-0.122*	-0.0701	-0.170	-0.0476	0.0212	0.0108	0.0186
	(0.0729)	(0.237)	(0.219)	(0.226)	(0.154)	(0.0549)	(0.163)
Ed High	-0.388***	-0.645**	-0.0684	-0.631**	-0.361*	-0.150**	-0.687***
	(0.0925)	(0.279)	(0.245)	(0.271)	(0.215)	(0.0745)	(0.232)
F Size Low	1.899***	1.978***	2.048***	1.877***	1.358***	0.864***	1.678***
	(0.0858)	(0.353)	(0.305)	(0.336)	(0.323)	(0.123)	(0.217)
F Size High	-4.343***	-3.848***	-4.140***	-3.854***	-1.157***	-0.928***	-4.589***
	(0.0689)	(0.213)	(0.203)	(0.204)	(0.144)	(0.0543)	(0.151)
Cons Low	-0.730***	-0.992***	-0.653***	-0.925***	-0.338**	-0.148**	-0.926***
	(0.0815)	(0.245)	(0.229)	(0.228)	(0.162)	(0.0599)	(0.182)
Cons High	0.0677	0.450*	0.457**	0.527**	0.252	0.116	0.499***
	(0.0795)	(0.256)	(0.232)	(0.244)	(0.189)	(0.0729)	(0.187)
Treat	-0.589***	-0.567**	-0.374*	-0.737***	-0.407***	-0.145***	-0.710***
	(0.0727)	(0.225)	(0.214)	(0.216)	(0.149)	(0.0547)	(0.180)
Constant	1.293***	1.146***	0.680***	1.340***	0.608***	0.370***	1.609***
	(0.0866)	(0.257)	(0.231)	(0.254)	(0.177)	(0.0670)	(0.203)
Observations	19,319	4,525	4,525	4,525	4,525	4,525	4,525
R-squared	0.249	0.188	0.124	0.234	0.145	0.122	0.233

Robust standard errors in parentheses

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

Figure 11

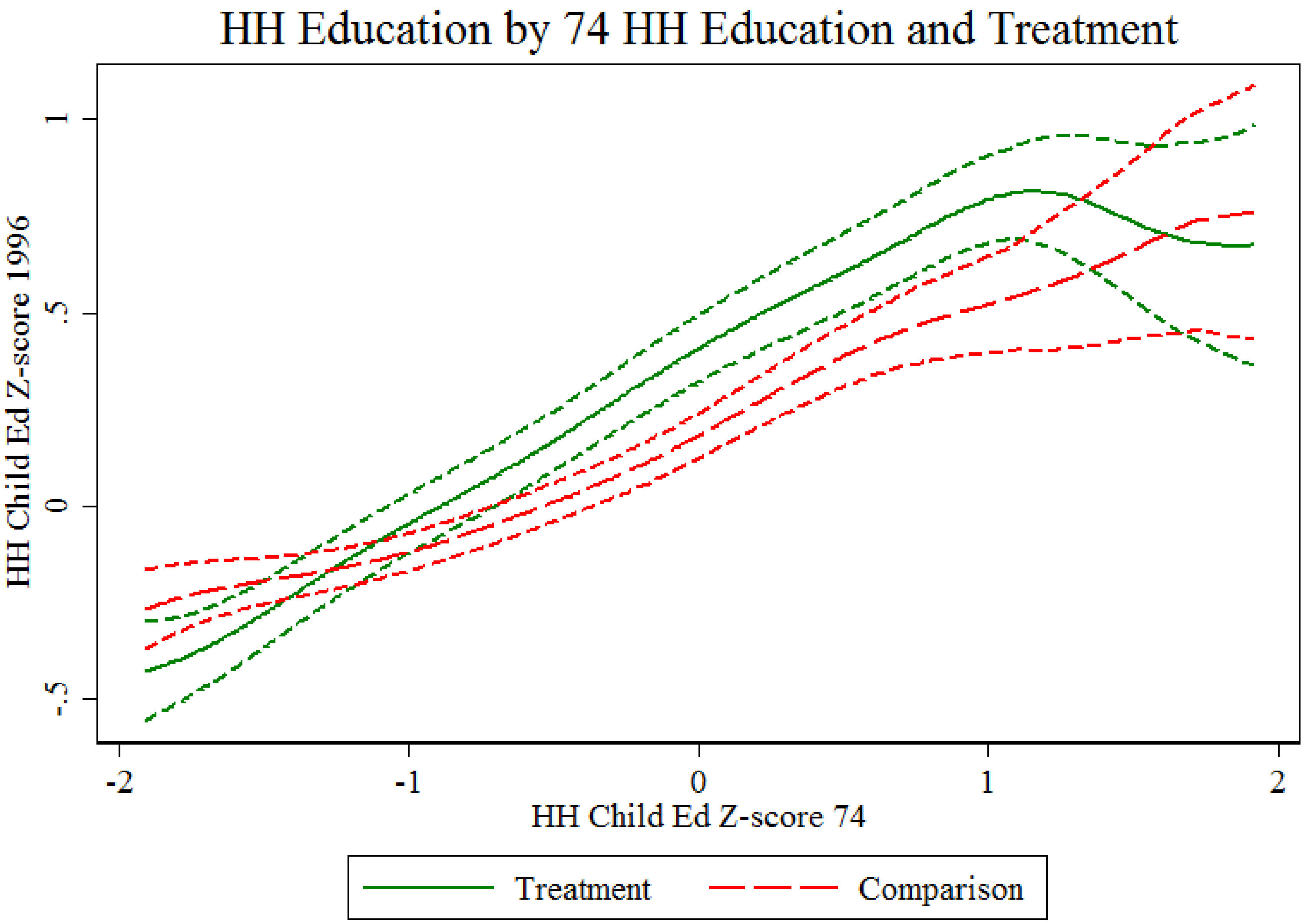

HH Education by 74 HH Education and Treatment
HH Child Ed Z-score 1996
-.5
0
.5
1
-2
-1
0
1
2
HH Child Ed Z-score 74
Treatment
Comparison

Table 6

HH Ed Z-score 96 by 74 Conditions and Treatment

VARIABLES	(1) Formal	(2) Predicted	(3) 74 Weights	(4) 74/96 Weights	(5) 96 Weights	(6) No Weights
Ed Low	-0.0289	-0.0744	-0.0978	0.00556	-0.0459***	-0.199***
	(0.0841)	(0.0611)	(0.0832)	(0.0675)	(0.0164)	(0.0409)
Ed High	0.364***	0.272***	0.373***	0.230***	0.108***	0.288***
	(0.0781)	(0.0688)	(0.0686)	(0.0632)	(0.0188)	(0.0394)
F Size Low	-0.0342	0.0322	-0.0288	-0.00878	-0.00235	-0.0509
	(0.0994)	(0.0937)	(0.0975)	(0.0905)	(0.0304)	(0.0564)
F Size High	0.0836	0.119**	0.0601	0.0212	0.0172	0.0747**
	(0.0659)	(0.0586)	(0.0576)	(0.0498)	(0.0146)	(0.0347)
Cons Low	-0.167***	-0.145**	-0.148**	-0.0599	-0.0418***	-0.142***
	(0.0645)	(0.0576)	(0.0608)	(0.0522)	(0.0134)	(0.0352)
Cons High	0.204***	0.199***	0.190***	0.116*	0.0769***	0.191***
	(0.0789)	(0.0730)	(0.0706)	(0.0628)	(0.0203)	(0.0423)
T x Ed Low	-0.391***	-0.376***	-0.303***	-0.194**	-0.0442*	-0.212***
	(0.119)	(0.0986)	(0.115)	(0.0984)	(0.0255)	(0.0621)
Treat	0.216***	0.242***	0.130**	0.0606	0.0459***	0.192***
	(0.0697)	(0.0642)	(0.0625)	(0.0522)	(0.0164)	(0.0410)
Constant	-0.131*	-0.207***	-0.0582	-0.0145	-0.0157	-0.0779
	(0.0777)	(0.0666)	(0.0761)	(0.0620)	(0.0185)	(0.0488)
Observations	3,322	3,322	3,322	3,322	3,322	3,322
R-squared	0.068	0.035	0.120	0.085	0.064	0.114

Robust standard errors in parentheses

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