

Date: March 7, 2004

Altruism, Household Coresidence and Women's Health
Investment in Rural Bangladesh

Andrew D. Foster
Departments of Economics and Community Health
Brown University

This research supported in part by NIH Grant HD10379.

I. Introduction

In recent years there has been increased analysis of the distribution of household resources both within households and across households that are linked by family ties. A common feature of this literature, which has addressed a variety of issues such as the extent to which income shocks are redistributed across households as might be expected in the presence of family-based altruism (Altonji, Kotlikoff, Hayashi 1992, 1997) and whether the distribution of resources within households is importantly influenced by the bargaining position of individuals or groups within the household (McElroy 1990; Thomas 1990; Schultz 1990; Hayashi 1995), is that it has not generally accounted for the fact that coresidence by related sub-families is a choice that may both influence and be influenced by the resource allocation decisions. Of particular concern is the fact that the selection process associated with the decision to coreside may bias statistical estimates leading to inappropriate inference.

In this paper, I posit a simple parametric model of family resource distribution and joint residence and use this model to examine the proposition that jointly-resident family units care differently about the allocation of resources to each other than do otherwise equivalent family units living in separate households. In particular, I examine the extent to which differences in measures of parental human capital result in differences in the nutritional status of adult women in the household.

The focus on women's health is motivated by the longstanding concern about the vulnerability of women in rural Bangladesh given social constraints on activity and mobility and the prominence of patrilocal residence (Cain 1985; Rahman, Menken and Foster 1992) The latter feature is of particular interest because it implies that co-resident married women are, for the most part, biologically unrelated to other household members; indeed, it has often been

argued that daughters-in-law face particularly difficult circumstances when joining their husband's household if the husband remains coresidence with his parents or siblings. The relationship between a daughter-in-law and her husband's mother may also influence the degree of support provided to the latter, although this possibility has received considerably less attention in the literature. In any case, it appears likely that an analysis of the nutritional status of women and its relationship to living arrangements is likely to provide new insight into the position of women within the household as well as to provide more general insights into behavioral differences between joint and non co-resident families.

Results indicate that unit-specific attributes have a limited influence on maternal outcomes net of a household-specific effects in a joint household while such effects are observed in recently divided households net of family-specific (across household) effects. While these results are consistent with the notion of greater altruism among coresident units, it is also established using the model that such a pattern might also arise from selectivity associated with the household residence decision if the nutritional status of coresident women from different family units are viewed as complementary. A simulation estimator that corrects for these selectivity effects is then developed for which identification comes from the characteristics of non-coresident units in the case of the within household estimator and from the distribution of family unit characteristics within households for the across-household estimator. Application of this procedure indicates that the selection effect, while changing coefficient estimates substantially, is not entirely responsible for the observed differences between the within household and across household human capital investment equations. A second procedure is constructed that makes use of body size measures at two points in time separated by 16 years. The results conform closely to those obtained using the cross-sectional approach.

II. Theory

The theoretical model is designed to yield an analytic basis for the notion that selection associated with joint residence may importantly bias estimates of health investment equations, and to provide a foundation for the development of an econometric procedure that adjusts for potential selectivity. Necessary features of this model are that it should endogenize both the household residence decision and health investment in joint and non-joint household families, that it should produce relatively transparent predictions, and that it should account for the possibility that resources may be transferred across households as well as redistributed within households. The model is related in spirit to that developed in Foster and Rosenzweig (1998), extended to incorporate health as a household-specific public good and to allow for unobserved taste differentials across family units.

A central feature of the model relates to how decision-making among distinct units and households is coordinated (i.e., who operates as the residual claimant, how reservation utilities are determined, and how any surplus is distributed). In particular, it is assumed that (1) that the choice of the allocation of units to households, consumption and human capital allocations for each unit, and a level of the public good for each household is efficient (2) the utility function exhibits transferable utility (see Bergstrom 1997). The first assumption seems a reasonable characterization of family behavior in that a group of related family units that live in close proximity to one another should have ample time to uncover and exploit existing opportunities to make one unit better off without decreasing the welfare of any other unit. The second assumption provides a substantial payoff in terms of analytical convenience but also imposes substantial structure on the model. The payoff is that given transferable utility all distributional considerations both within and across households can be satisfied through the direct transfer of

consumption c_i and thus distributional concerns do not affect the efficient choices of household residence patterns, human capital investment, and the level of the public good. The cost is of course that the implication that human capital investment is not affected by distributional concerns may not be, in fact, reasonable.

A general characterization of the model is as follows. Consider a family that consists of I distinct family units partitioned into K households. It is assumed that the head of each unit i has preferences defined over private-consumption c_i , the level of health in the units family h_i , the average health within the household \bar{h}_{k_i} where k_i denotes the household of unit i , and a household-specific public good z_{k_i} :

$$u(c_i, h_i, \bar{h}_{k_i}, z_{k_i}; v_i) \quad (1)$$

where v_i is a unit-specific parameter reflecting intrinsic demand or taste for health for that unit.^{1 2}

Intra-family transfers across households are freely permitted so that the relevant budget constraint equates income and expenses at the level of the family:

$$\sum_{i=1}^I c_i + p_h h_i + \sum_{k=1}^K z_k = \sum_{i=1}^I y_i \quad (2)$$

¹Arguably one might also want to include, in addition to household-specific average schooling, the family-specific average health. Inclusion of this argument has a small effect on the interpretation of parameter estimates but does not alter the econometric approach. Thus for notational simplicity this possibility has been ignored.

²I use the term intrinsic demand to reflect the idea that this parameter is an indicator of the demand for health in the family unit if it were autarchic and thus could, in principle, be derived from an underlying model of health demand under autarchy. In this paper, however, I will treat this component of health demand as a reduced form so that I can focus on the implications for health of combining family units.

Under these assumptions the family's optimization problem involves the choice of the household allocations k_i , health investment h_i , and consumption levels c_i for each unit i and of the level of public good z_k in each household k to maximize

$$W = \sum_{i=1}^I u(c_i, h_i, \bar{h}_{k_i}, z_{k_i}; v_i) \quad (3)$$

subject to (2). The health investment equation for family unit i that arise from this optimization problem then has as arguments total family income, prices, and the intrinsic demands of all units.

In order to use this model for econometric analysis it is helpful at this point to impose functional forms. In particular, it is assumed that the utility function is quadratic in its arguments with three exceptions. First, consistent with the assumption of transferable utility there are only linear terms in c . Because transferable utility would in general permit the marginal utility of consumption to be a function of public goods one could in principle permit interactions with household-level health for the purpose of examining behavior in joint households; however, this would lead to violations of transferable utility when considering the behavior of families living in separate households, which would have different average levels of health. Second, it is assumed that the intrinsic demand enters into the utility function in such a way that it appears as a linear term in the marginal utility of own-family health. This assumption makes the parameter readily interpretable as a measure of the taste for own health. Third, it is assumed that utility is strongly separable in z . This requirement also imposes significant structure, but decreases the complexity of the problem yielding more transparent theory and simplified econometric

analysis.³ Given these assumptions the utility function is:

$$u(c, h, \bar{h}, z; v) = a_0 c + (a_1 + v)h + a_2 \bar{h} + a_3 z - \frac{1}{2} h^2 - \frac{a_5}{2} \bar{h}^2 - \frac{a_6}{2} z^2 + \frac{a_7 + a_5}{2} h \bar{h} \quad (4)$$

where I have specified the parameters and incorporated an arbitrary normalization in such a way as to obtain relatively clean analytical expressions for the human capital investment equations below.

Given this functional form it is relatively simple to show that the relative magnitude of the relationship between intrinsic demands and health investment that arise when differencing across or within households is critically related to the extent to which units care about the average health in their respective households. Consider first the case of a two-unit family. In this case health investment for unit i is

$$h_i = \frac{1}{1 - a_7} (a_1 + a_2 - a_0 p_h + v_i) \quad (5)$$

in the case of separate residence and

$$h_i = \frac{1}{1 - a_7} (a_1 + a_2 - a_0 p_h) + v_i + \frac{a_7}{1 - a_7} \bar{v} \quad (6)$$

³Given this restriction and the previous one it is shown below that only the difference in the v across and within households affects the returns to joint residence. Without this restriction the levels of v would also affect these returns because they would then influence the demand for z . The primary econometric implication of relaxing this restriction would be that one could not permit, without further assumptions, family heterogeneity in intrinsic demands that is correlated with unit-specific observables. If these effects were assumed to be absent then it would be straightforward to test the implied exclusion restriction.

for joint residence where \bar{v} is the average intrinsic demand for the two families. Note that the average intrinsic demand only appears when units are joint residence, reflecting the fact that one unit only cares about the health of the other when they coreside.⁴ Differencing these two equations across and within households, respectively, yields

$$h_2 - h_1 = \frac{v_2 - v_1}{1 - a_7} \quad (7)$$

and

$$h_2 - h_1 = v_2 - v_1 \quad (8)$$

respectively. Because $a_7/4$ may be shown to be cross partial of the utility function with respect h_1 and h_2 with $(h_1 + h_2)/2$, $a_7 > 0$ implies that own h and the h of the other unit are complementary (assuming $a_7 < 1$, which is required to ensure a local maximum). If this condition is met then a given within-family difference in intrinsic demand maps into a larger difference in health in the absence of coresidence than given coresidence. Conversely $a_7 < 0$ implies that own and cross-values of h are substitutes and thus differences in health are larger under joint residence.

The equation dictating whether joint or separate residence is preferred is also relatively straightforward and indicate that the effects of variation in intrinsic demand for health on joint residence depends again on whether own and cross-values of health within the household are substitutable or complementary. Let $W(\mathbf{k})$ denote residence-pattern conditional total welfare (i.e., the maximum of (3) over resource allocations for given residence) where \mathbf{k} is the vector of

⁴If average family health were also an argument of the utility function then the average family intrinsic demand would also appear in the former equation.

household indicators so that $\mathbf{k}=[1,2]$ denotes separate residence and $\mathbf{k}=[1,1]$ denotes separate residence. Then

$$\mathcal{W}([1,2]) - \mathcal{W}([1,1]) = \frac{1}{4} \frac{a_7}{1-a_7} (v_1 - v_2)^2 - \frac{a_3 a_0 p_z}{a_6} + \frac{3}{4} \frac{a_0^2 p_z^2}{a_6} \quad (9)$$

Note that if $a_7 > 0$ (complements) the first term is negative. Thus if own and cross health are complementary then household division is more likely the higher the difference in intrinsic demands while the opposite results holds if they are substitutes.

Extending the model to three units is important because, as will be argued below, identification of the model cannot easily be achieved in two-unit households. Fortunately, although this somewhat complicates the analysis because of the increased number of potential residence patterns that must be considered, it does not alter fundamentally the nature of the problem. Indeed, although the level human capital equations become more complicated the predictions regarding within household differences (equation 7) are identical and the equations regarding across household differences (equation 8) are identical if household averages are computed. For example, if units 1 and 2 are allocated to the same household and 3 is in a separate household then

$$h_2 - h_1 = v_2 - v_1 \quad (10)$$

and

$$h_3 - \frac{h_2 + h_1}{2} = \frac{1}{1-a_7} \left(v_3 - \frac{v_2 + v_1}{2} \right) \quad (11)$$

The welfare comparisons across residence patterns given a three-unit family are also readily interpretable. Consider first different ways of dividing up the three units into two households. Comparison of the case in which 1 lives with 3 with that in which 1 lives with 2 yields

$$\mathcal{W}([1,2,1]) - \mathcal{W}([1,1,2]) = \frac{1}{2} \frac{a_7}{1-a_7} (u_3 - u_2) \left(\frac{u_2 + u_3}{2} - u_1 \right) = \frac{1}{4} \frac{a_7}{1-a_7} ((u_1 - u_2)^2 - (u_1 - u_3)^2) \quad (12)$$

The latter form of the expression, for $a_7 > 0$, implies that coresidence of 1 and 3 is preferred to coresidence of 1 and 2 if and only if the intrinsic demands of 1 and 3 are more similar than are the intrinsic demands of 1 and 2. Comparison of the case of separate residence of all three units with that in which 1 lives with 2 yields the RHS of equation (9):

$$\mathcal{W}([1,2,3]) - \mathcal{W}([1,1,2]) = \frac{1}{4} \frac{a_7}{1-a_7} (v_1 - v_2)^2 - \frac{a_3 a_0 p_z}{a_6} + \frac{3}{4} \frac{a_0^2 p_z^2}{a_6} \quad (13)$$

Finally comparison of the case of joint residence of the whole family and coresidence of 1 and 2 yields

$$\mathcal{W}([1,1,1]) - \mathcal{W}([1,1,2]) = -\frac{1}{6} \frac{a_7}{1-a_7} \left(v_3 - \frac{v_1 + v_2}{2} \right)^2 + \frac{a_3 a_0 p_z}{a_6} - \frac{7}{12} \frac{a_0^2 p_z^2}{a_6} \quad (14)$$

Thus joint residence is preferred to this particular two-household alternative for complementary health if the intrinsic demands in the third unit do not differ substantially from the average intrinsic demand in the first and second units.

In order to use this model to better understand the likely sorts of bias that arise when household residence is endogenous with respect to health outcomes it is necessary to posit a relationship between the v_i , which have heretofore been considered exogenous measures of intrinsic demand for health, and observable data. In particular, I assume that these intrinsic demands can be written as a linear function of a vector observed characteristics x_i for each individual in the household, a fixed family-specific effect η that may be correlated with the x_i , and an orthogonal individual-level error term:

$$v_i = x_i \beta + \eta + \epsilon_i \quad (15)$$

A key aspect of this model that is evident in this specification is that v_i cannot depend on household residence patterns or consequences of household residence patterns such as total household assets.⁵ Note also that the effects of family-specific observables as well as the family-specific unobservable η cannot be identified from equations characterizing residence choice and differenced human capital investment given the assumptions of the model.

III. Empirical Implementation

Substitution of (15) into the equations for differenced health investment and differences in residence-conditional family welfare provides an illustration of the bias introduced by the endogeneity of joint residence. In particular, equations (10) and (11) become

$$\Delta h_i = \Delta x_i \beta + \Delta \epsilon_i \quad (16)$$

⁵I do not at this point need to take a position on whether the β may be interpreted structurally. If there is unobserved heterogeneity that influences intrinsic demands for human capital and is correlated with the x 's then the β s may be thought of as a result of projecting these demands on the set of observed variables.

and

$$\Delta \bar{h}_k = \frac{1}{1-a_7} (\Delta \bar{x}_k \beta + \Delta \bar{\epsilon}_k) \quad (17)$$

with Δ in (16) and (17) differencing within and across households in the same family, respectively. The bias arises from the fact that the conditional independence of ϵ given x in (15) is not retained when one examines data in which residence is chosen optimally. In particular, the implications of the model are, as noted, depending on the structure of preferences. Assume, for the moment that own and cross-health are complementary. Then the model implies that units are most likely to coreside when they have similar intrinsic demands for schooling. Thus, for example, if two units have a large positive observable intrinsic demand difference $\Delta x_i \beta$ and they nonetheless choose to coreside then one may infer that they have a compensating difference in the unobservable component of intrinsic demands $\Delta \epsilon$, yielding $\sigma(\Delta x_i \beta, \Delta \epsilon_i) < 0$. Conversely, if two units have a large positive observable intrinsic demand difference $\Delta x_i \beta$ and they live separately then one may infer that they do not have a compensating difference in the unobservable component of intrinsic demand $\Delta \epsilon$, yielding $\sigma(\Delta \bar{x}_i \beta, \Delta \bar{\epsilon}_i) > 0$. Thus β estimates will tend to be biased toward zero for within household estimation and $\beta/(1-a_7)$ estimates biased away from zero for across household estimation. Opposite effects will be observed if $a_7 < 0$: units with large differences in $\Delta x_i \beta$ who live separately will not have compensating differences in the unobserved component of intrinsic demand.

Given that the source of bias is joint residence and joint residence is determined, given the model, based on both the observable and unobservable components of y , instruments are unlikely to be available to construct a consistent instrument-variable estimator: anything that is

correlated with differences in the x variables will also tend to be correlated with the ϵ . A more promising approach is to treat this problem as one of selectivity bias, that is one in which the set of households that are in fact observed are a non-random sample of all potential households. Again, however, standard procedures do not easily apply because the nature of the relationship between unobservables determining the selection process and those determining the human capital process is not a simple linear one.

I therefore adapt the two-stage simulation approach of Foster (1997), which addresses selection bias arising from the process of matching in the marriage market provides. The procedure has two stages. During the first stage, a simulated method of moments procedure applied to individual-level longitudinal data on residence patterns is used to estimate the β and other parameters up to an arbitrary scalar.⁶ In particular, let $I_j([1,2])$ denote the event that household j chooses separate residence. Then equation (9) implies that

$$E(I_j([1,2]) - \text{Prob}\left(\frac{a_7}{1+a_7}(\Delta x\beta + \Delta\epsilon)^2 + \tau_{12} > 0\right) | x_2, x_1) = 0 \quad (18)$$

where τ_{12} is a household-size specific based on the second term in (9).⁷ While in the case of family sizes of two this expression has a straightforward analytic solution for certain distributional assumptions such as the normal, the same cannot be said for larger families. I therefore adopt a simulation procedure in which the ϵ_i are drawn from a particular distribution (the standard normal) and then the value of the expression inside the $\text{Prob}()$ function of (18), for

⁶As is typically the case for limited-dependent variable models the variance of ϵ cannot be separately estimated from the regressor coefficients.

⁷Note that the term $a_7/(4(1-a_7))$ has been absorbed into τ_{12} as it cannot separately be identified using only residence data.

given β and τ_{12} may be computed. Across many simulations this approach yields consistent estimates of the second term in (18). In practice, however, it has been established (Pakes and Pollard 1989) that consistency may be obtained for this type of problem with respect to the number of families, with the number of simulations per family held fixed. I therefore carry out the analysis with just two draws on the ϵ per family. I also adopt the standard procedure of using extreme value errors appropriately scaled to smooth the objective function to aid computation. Thus, the actual moment condition for two-unit family is

$$E(I_j([1,2]) - \frac{1}{1 + \exp(-(\frac{a_7}{1+a_7}(\Delta x \beta + \Delta \epsilon)^2 + \tau_{12})/\delta)} | x_2, x_1) = 0 \quad (19)$$

for small δ . For three-unit families there are 5 distinct household allocations, as noted, so four distinct moment conditions must be constructed. For example, the moment condition for the probability that units 1 and 2 coreside and 3 is in a separate household may be written using the expressions in equations (12)-(14) as

$$E(I_j([1,1,2]) - 1 + [1 + \exp((W([1,1,1]) - W([1,1,2]))/\delta) + \exp((W([1,2,3]) - W([1,1,2]))/\delta) + \exp((W([1,2,1]) - W([1,1,2]))/\delta) + \exp((W([1,2,2]) - W([1,1,2]))/\delta)] = 0 \quad (20)$$

The simulation of household residuals may also be used to construct consistent estimates of the human capital investment equations, given estimates of β obtained from the household residence equations. The approach rests on the recognition that equation (16) and (17) could be estimated consistently if the ϵ were actually observed. Let W denote a matrix consisting of the Δx_i and $\Delta \epsilon_i$ and let H denote the vector of $\Delta \epsilon_i$. Then consistent estimate of the column vector $B = [\beta \ 1]^T$ may be obtained using least squares

$$\hat{B} = W'W^{-1}W'h \quad (21)$$

Alternatively, if there existed a vector Z of variables that predict $\Delta\epsilon_i$ given Δx_i , a less efficient instrumental variables procedure may be used

$$\hat{B}^{IV} = (W'Z(Z'Z)^{-1}Z'W)^{-1}(W'Z(Z'Z)^{-1}Z'H) \quad (22)$$

While given the unobservability of the ϵ_i , $W'Z(Z'Z)^{-1}$ cannot of course be computed directly, a consistent estimate may be obtained by computing this expression with a data set on household residence using the actual x_i but then simulating the ϵ_i and then using the above expressions inclusive of the estimates of the β and τ coefficients obtained from the first-stage of estimation.

Appropriate instruments for the within-household estimator are derived from the predicted values of $x_i\beta$ of the excluded unit or units. The intuition is straightforward. As noted the model predicts that three-unit households will be sorted so that those with the most similar values of v_i will reside jointly. Suppose $x_1\beta > x_2\beta > x_3\beta$ and that units 1 and 2 coreside. Then the closer $x_3\beta$ gets to $x_2\beta$, the more negative would have to be $\epsilon_1 - \epsilon_2$ in order for this choice to be optimal. Moreover, $x_3\beta$ is excluded from the differenced regression (16). Thus levels of $x_3\beta$ and second and higher-order interactions of $x_3\beta$ with $x_1\beta$ and $x_2\beta$ can serve as instruments for the estimation of (16).⁸ An analogous argument applies to the across estimator, (17), but in this case the appropriate instruments involve the variance of within-household observable indices. In particular, if $x_1\beta > x_2\beta > x_3\beta$, units 1 and 2 coreside in household 1, and unit 3 is in household 2

⁸Note that the assumption of an ordering is not innocuous. Without such an assumption and if the ordering of units within the family were random, $x_3\beta$ would not be correlated with $x_2\beta - x_1\beta$. But of course as the $x_i\beta$ are observable, units may be ordered in any desirable way. Or, interactions between the differenced $x_1\beta$ and $x_3\beta$ may be incorporated into the instrument set.

then the smaller is $\bar{x}_1\beta - \bar{x}_2\beta$ the larger must be $\bar{\epsilon}_1 - \bar{\epsilon}_2$ in order that residence in two household is preferred to residence in one household.

Cross-Sectional Data

The primary data for this analysis comes from 1997 Matlab Health and Socioeconomic Survey (MHSS), which was designed by the author and three collaborators (Rahman *et al* 1997). These data contain detailed information on household composition, assets and other household resources, and individual-level measures of adult and child human capital inclusive of schooling, body size, physical performance and cognitive skills on 4539 households in the Matlab study area of the International Centre for Diarrhoeal Disease Research in Bangladesh (ICDDRDB). The data have been linked using individual registration numbers to the vital and census records for the that Matlab study area providing a longitudinal record on household formation and dissolution from the mid 1960s until the time of the MHSS. The data are well suited to the issues addressed in this paper in that the longitudinal vital records permit tracking of family units over time, as required for the estimation of the household residence model, and the MHSS sample is designed to include non-coresident family units, as required to consistently estimate the specified across and within-household differenced health investment equations.

In order to ensure that any differences in the nutritional status of coresident and non-coresident married do not are from differences in residence (and thus, for example, differences in the prices faced) the analysis focuses on non-co-resident family units living in close proximity. Specifically, the data employed include information on multiple households within the same compound or “bari”. This social unit, which is a prominent feature of life in that country, may be characterized as “...a cluster of households usually around a common yard accommodating households whose heads are related by blood or affinal connections” (Caldwell 1982: 36). Most

young men in rural areas that no longer coreside with parents or siblings can be found in the same bari as these other members of their family (Foster (1993)).

The bari in Bangladesh has been argued to play an important role as a source of support. Makhlisur Rahman (1986), for example, notes that there is “a marked reluctance on the part of wealthy individuals...to permit (fellow) bari members to sink below a certain minimum level”. It is not clear, however, how the extent of resource-sharing among households and, in particular, the relative well-being of married women in different households within the bari, compares with that that might be expected by coresident family units. As might be anticipated if there are important differences in the degree of mutual support provided among those in the same bari and those in the same household, the process of household partition, in which two family-units become separate households, is a prominent event in South Asia. As Caldwell (1984 p128) reports, household partition is “...one of the central facts of social life, observed and discussed by all. There is little in the way of an intermediate or blurred situation: when partition occurs, division, somewhat ceremoniously and usually without rancour, is affected in eating arrangements, the family budget, land (if any), and residential arrangements”. Indeed, the limits to within-bari support are also evident in Rahman’s characterization of the bari as one in which a “minimum-level of support” is provided. The only empirical evidence on this question, of which I am aware, is that presented in Foster (1993), which reports evidence consistent with the notion that the bari provides less complete pooling of resources than does the household, but this evidence is subject to precisely the kind of selection bias discussed above.

Data preparation for this analysis involved two stages. The first stage was to select from the MHSS information on ever married women. The number and sex composition of coresident children and the education and age of a husband, if present, were then linked to the women

record. The second stage was to determine the group of women that could reasonably be linked as family units that could have been coresident. This grouping was done with close attention to the notion of the bari and the fact of patrilocal residence. In particular, all women living in a given bari in 1996 at the time of the MHSS who were ever married and coresident with each other in 1982 were grouped together. In addition, women who had married since 1982 were grouped with other women from the same bari if their spouse was coresident with other members of the group or their spouses in 1982. Groups with more than three women made up only 12% of all baris. These large groups were excluded for reasons of tractability and because typically only two households participated in the MHSS from any one bari. These procedure resulted in the selection of 1850 groups, 33% of which were one-unit families, 46% of which were two-unit families, and 21% of which were three-unit families. Among the three-unit families a majority coresided in one household, but there was substantial representation of other configurations as well.

Cross-sectional Results

The above model, as noted, is sufficiently general to capture two quite different sets of circumstances, based on the extent to which the health of members of other family units are viewed as substitutable or complementary to own health. In the case that health is substitutable, households tend to be formed of people who are different from each other in terms of their underlying intrinsic demand for health, while if health is complementary across family units then households tend to be formed of units with similar intrinsic demands for health. Because these indices are posited to be linear functions of observable attributes, one might expect to be able to distinguish these cases by examining within and across variation in observables, recognizing that this is a somewhat imperfect test because of the presence of unobservable, correlations among

variables, and differences in the extent to which different variables affect intrinsic demands.

Measures of the standard deviations within and across households for the key variables of interest are presented in the first two columns of Table 2. For the most part these estimates suggest that the within variance is higher than the across variance. The largest difference is observed for women's education where the within household standard deviation is 28% higher than the across standard deviation. There are three cases, where the relative magnitudes are reversed: age, husband's education, and the number of sons, although in no case is the across standard deviation more than 5% higher than the within.

The results from the structural household division estimation are presented in Table 2. Because of the inability to identify the magnitude of the parameter a_7 directly, it is helpful to assign *a priori* a sign for a_7 , which captures the degree of complementarity, before carrying out the estimation, thus yielding two sets of results. While the moment conditions are satisfied for both sets of criteria, the model with complementarity is a better fit to the data because it yields a lower mean-squared error.⁹ Moreover, by using the relevant coefficient estimates to simulate out household division it becomes clear that the use of the criteria of comparing standard deviations of individual attributes in Table 1 can be quite misleading. Columns 3-6 of Table 1, which present within and across standard deviations based on the simulated residence given the two sets of coefficient estimates in Table 2, show that both conditions generally yield higher within than across standard deviations for specific household attributes. While this may seem counterintuitive it should be noted that both approaches attempt to fit the available data which

⁹Because a small number of simulations per family are drawn the mean squared error is estimated using the product of the prediction error of two independent simulations for each family. It may be shown that this yields an expression with the same probability limit as would be obtained with a large number of simulations per family.

exhibit this pattern, and that when the index $x\beta$ is computed based on the respective set of coefficients the expected pattern is observed: the within standard deviation of this index is less than the across standard deviation given the complementary assumption while the opposite result holds under the substitutability assumption. Given these results, subsequent discussion will focus on the complementarity results, unless otherwise stated.

The coefficient estimates of the household division model in Table 2 , as noted, reflect the effects of the covariates on intrinsic demand for health investment, which in combination with the intrinsic demands of other family units, determine household residence choices. Unfortunately the direction of the effects cannot be inferred from these coefficients given the structure of the model. Instead, significant coefficients indicate, for the model with complementarity, that this variable is one for which variability predicts a lower likelihood of joint residence. Thus, for example, the coefficient on husband's education indicates that household are unlikely to coreside if household members have substantially different levels of schooling. Similarly, there is evidence that households with different numbers of sons and daughters do not tend to coreside. Both the husband's age and wife's age exhibit strong quadratic effects, although over the relevant range the effects of husband's and wife's age exhibit opposite signs.

The effects of the coefficients are moderate in size. Assume that there is a single-year of schooling differential between two ever married women in a given family. Then a unit additional increase in schooling for the more schooled individual decreases the probability of joint residence by 4.1 points. If one woman has one more son than the other, then an additional son for the first woman results in a 4.7 point decrease in joint residence. One the larger effects is associated with whether a husband is present. If there is a one unit difference in the predicted

index for two women with the same marital status, then the change in whether a husband is present(e.g., through the migration of a spouse) results in a 17 point decline in the probability of joint residence.

Table 3 presents the estimates of the health investment equations. The dependent variable is maternal BMI and the right-hand side variables are the same attributes as used in the family residence decision. Note that although the model incorporates the assumption that these coefficients should be proportional to those obtained from the structural analysis, this restriction need not hold if BMI imperfectly measures health investment in the sense that the X variables influence BMI net of health investment as determined in the model. This might be the case, for example, because BMI does not adjust for age although there may be important effects of aging on body size that are not themselves directly related to the provision of resources to the person in question. Under such circumstances the coefficient on age in the health investment equation relative to that, say, for education, would be different than the same ratio taken from the residence-equation estimates.

The first and third column contain estimates of the differenced health investment equations where differencing is carried out within and across households, respectively. While the precision of the estimates is not large, the overall impression that emerges is that the coefficients for the across estimator are somewhat larger and more likely to be significantly different from zero than are the within estimates. As noted, this might be evidence of complementarity model, it could arise either because of there are real differences in the effects of family-unit attributes on health within and across households, as posited by the model, or because of selectivity bias.

Selectivity correction for the within household does not substantially change the results, a

result that may in part be explained by the relative weakness of the excluded instruments in predicting the errors given family residence ($P=.31$). On the other hand, the selectivity corrected estimates for the across estimates, for which the excluded instruments in the first stage were significant at the 7% level, differ substantively from those obtained without this correction. In particular, the coefficients on key husband's attributes including whether the women is currently married, whether her husband is absent (likely reflecting earnings of a husband with employment in the city), and husband's schooling increase significantly. The husband's schooling estimates, for example, suggest that a 10 percent in husband's education at the mean level of 3 years results in a .3 percent increase in maternal body size. Being currently married instead of being widowed or divorced results in a 11% increase in body size, a substantially larger figure than would have been obtained using the uncorrected estimates. Finally, daughters-in-law appear to be disadvantaged in terms of intrinsic demand for health—the coefficient is negative for both the split and intact households once selection is accounted for, and it is significant for the split households. The sign of the bias is also as expected—bias is away from zero for the split households and toward zero in the intact households.

Longitudinal Results

An alternative approach to addressing the consequences of household division is to consider changes over time. It is argued that by observing the same individual at multiple points in time that fixed attributes of the individual that are correlated with the possibility of joint residence are thus controlled and that the problem of endogenous household residence is thus substantially diminished. It is clear in the context of the above theoretical framework that this argument may fail for two reasons. The model clearly indicates that both individual outcomes and household residence depend not only on the characteristics of the individual but also in the

distribution of characteristics in the family.

To address the possibility of change over time in intrinsic demand for household public goods I add a time sub-script to the β in (15)

$$v_{it} = x_i \beta_t + \mu + \epsilon_{it} \quad (23)$$

The change in health for the member of an intact two-unit family is thus

$$\Delta_t h_{it} = -\frac{a_0}{1-a_7} \Delta p_{it} + \left(x_i + \frac{a_7}{1-a_7} \bar{x}\right) \Delta_t \beta_t + \Delta \epsilon_{it} + \frac{a_7}{1-a_7} \Delta_t \bar{\epsilon}_t \quad (24)$$

while the change in health for the member of a two-unit family that divides is

$$\Delta_t h_{it} = -\frac{a_0}{1-a_7} \Delta p_{it} + \frac{a_7}{1-a_7} \beta_t x_i + \frac{a_7}{1-a_7} (x_i - \bar{x}) \Delta \beta_t + \frac{a_7}{1-a_7} (e_{it} - \bar{e}_t) + \frac{1}{1-a_7} \Delta_t e_{it} \quad (25)$$

Notice that in the absence of changes in the ϵ_{it} over time there is no bias from estimating (24) using least squares. Moreover, even with such change it is not *a priori* clear what the sign of the bias will be given these changes. If intrinsic demands for unit i increase then this may increase or decrease probabilities of joint residence depending on i 's initial intrinsic demand relative to that of the other members of the household.

In contrast, for split households, bias arises even in the absence of changes in the shock. The reason is that household composition changes over time and thus the unobservables at two points in time enter in a different way as a result of the fact that the allocation across family units

will be different under coresidence. In this case, if both the difference in the x_i between i and the household average and the corresponding difference in the ϵ_i are both positive for positive β then a household split is likely and there is thus a positive correlation induced between the observable and unobservable components of the difference between own and average effects in (25). To address this problem I again make use of simulated errors—intuitively, if there is a great deal of variation in the unobservables across family units at time t and the household nonetheless stays together then it is likely that the residual difference compensates for the

To implement the analysis I use the Determinants of Natural Fertility Survey (DNFS) sub-sample of the MHSS. The DNFS was a survey of 2441 women 11-50 in January 1975 from 14 of the Matlab Study villages. These women were included in the MHSS framework and thus permit the analysis of body size change over a period of 21 years, a sufficiently lengthy period for a substantial degree of household division to have taken place. Of the original 2441 women, I identified 1213 who lived in multiple-unit households in 1975 and were reinterviewed as part of the MHSS in 1996.

For the purpose of this analysis, household family units were constructed based on the DNFS. Then women from the coresident family units were followed through 1990 using the MHSS records where possible and the Matlab Demographic and Surveillance System records otherwise. Body size data at two points in time were only available for primary respondents in the DNFS who were also interviewed in the MHSS. For other women coresident with these respondents in 1975 one could construct measures of demographic characteristics but not outcomes. Thus I cannot estimate the impact of household coresidence by differencing across time or across individuals at a point in time as above, but not both.

Results appear in Table 4, in which a pared down specification is used. Other variables

presented in Table 3 were not significant and thus are excluded so that the principle results might be highlighted—this is not surprising given that one is differencing over time for the same individual. Particularly in the case of an intact household, it is only possible to identify changes in coefficients over time.

Of particular interest are the coefficients in the fourth column, which correspond to split households. Comparison to (25) indicates that the coefficients on the individual variables in this specification indicate changes in the effects of the coefficients over time. Thus the coefficient on whether one's spouse was the son of the head in 1982 is negative and significant, indicating that intrinsic demand for health of daughters in law in the family (whether coresident or not) declined over this period. Given that the coefficient on the difference in this variable between the individual and family average is not different from zero, this result suggests that in fact daughters in law are allocated fewer resources than other women. This indeed seems consistent with the results from Table 3, in which the corresponding cross-sectional coefficient is negative. This result corresponding with the finding that body size is complementary within households suggests that the process of household division lowers body size for these women relative to others, particular the spouse and mother of the original head. By contrast educated women tend to have higher intrinsic demands for health, with this difference rising over time. Thus given redistributive effects of joint residence, educated women are better off when separately resident as one might expect.

Interestingly the magnitude of the coefficient on daughters-in-law becomes more negative once household division selection is controlled for. Evidently, daughters-in-law who split off from the original household are less disadvantaged than those in the intact households so the daughter-in-law effect is biased toward zero. This direction of the effect is as expected if bias

from the levels of the residuals dominates that from the change over time. To understand this note that I argued above that $x_i - \bar{x}$ is positively correlated with $\epsilon_{it} - \bar{\epsilon}_t$ in this sample. Thus the coefficient on the former is biased downward, which is what is observed. But one would anticipate that x_i is positively correlated with $x_i - \bar{x}$ so this induces a positive bias on the x_i term. Bias from the level differences is not present for the intact households so it is not inconsistent to see that the selection term is marginally significant and the bias operates in the opposite way in that case. For these households bias can only arise due to change in the ϵ .

Conclusions

In this paper I have examined the determinants of the nutritional status of ever married women in rural Bangladesh, giving particular attention to potential selectivity bias arising from the fact that joint residence is potentially endogenous with respect to health investment. In addition to examining the extent to which a woman's health is sensitive to the characteristics of woman's spouse as well as the number and composition of her children, the results generally support the notion that individuals care differentially about the welfare of coresident and non-coresident family members. Specifically, it appears that the health status of coresident family units are viewed as complementary to the health of one's own family. The results also suggest that intrinsic demand for health is lower for daughters in law than other female members of the household. Evidently joint residence is protective for these women given the complementarity of health allocations in the joint household.

References

- Altonji, Joseph G., Hayashi, Fumio and Laurence Kotlikoff (1992), "Is the Extended Family Altruistically Linked? Direct Evidence Using Micro Data," *American Economic Review* 82, 1177-1198.
- Altonji, Joseph G., Hayashi, Fumio and Laurence Kotlikoff (1997) "Parental Altruism and Inter Vivos Transfers: Theory and Evidence." *Journal of Political Economy*, 105(6):1121-1166.
- Aziz, K.M.A., (1979) "Kinship in Bangladesh", ICDDRDB Mongraph Series No. 1, Dhaka: ICDDRDB.
- Barkat-e-Khuda (1985), "The nuclearization of the joint family household in a rural area of Bangladesh", *Journal of Comparative Family Studies*, 16(3).
- Bergstrom, T. (1996) "A Survey of Theories of the Family", Mark Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, New York: Elsevier.
- Caldwell, J.C., G. Immerwahr, and L.T. Ruzicka (1982), "Illustrative Analysis: Family structure and fertility", *WFS Scientific Reports*, No. 39, Voorburg: International Statistical Institute.
- Caldwell, J.C., P.H. Reddy, and P. Caldwell, (1986) "The determinants of family structure in rural South India", *Journal of Marriage and the Family*, 46(1):
- Cain, M. (1985) "The fate of the elderly in South Asia" Implications for fertility", IUSSP International Population Conference, Florence, 4:279-91/
- Foster, A. D. (1993) "Household Partition in Rural Bangladesh", *Population Studies*, 47(1): 97-114, March.

- Foster, A.D. (1997) "Marriage-Market Selection in Rural Bangladesh", manuscript.
- Foster, A. D. and Mark R. Rosenzweig (1998) Household public goods, household division, and rural economic growth", manuscript.
- Foster, A. D. and Mark Rosenzweig, forthcoming, "Financial Intermediation, Transfers and Commitment: Do Banks Crowd Out Private Insurance Arrangements in Low-Income Rural Areas?" Andrew Mason and Georges Tapinos, eds., *Sharing the Wealth: Demographic Change and Economic Transfers between Generations*, Oxford: Oxford University Press.
- Hayashi, Fumio, (1995) Is the Japanese Extended Family Altruistically Linked? A Test Based on Engel Curves, *Journal of Political Economy*;103(3): 661-74..
- Heckman, J. (1979) "Sample Selection Bias as a Specification Error", *Econometrica*, 47 : 153-161.
- McElroy, M. (1990) "The Empirical Content of Nash-Bargained Household Behavior", *Journal of Human Resources*, 25: 559-83.
- Pakes, Ariel and David Pollard (1989) "Simulation and the asymptotics of optimization estimators", *Econometrica* 57:1027-57.
- Rahman, Makhlisur, (1986), "Tradition, development and the individual: A study of conflicts and supports to family planning in rural Bangladesh", eds., Penny Kane and Lado Ruzicka, *Asian Population Changer Series No. 1*. Department of Demography, ANU
- Rahman, M. Omar, A. Foster, and J. Menken, (1992) "Older widow mortality in rural Bangladesh", *Social Science and Medicine*, 34:1:89-96.
- Rahman,, O. A.D. Foster, P. Gertler, and J. Menken, (1998), Matlab Health and Socioeconomic Survey.

Schultz, T.P. (1990) "Testing the Neoclassical Model of Family Labor Supply and Fertility,"
Journal of Human Resources, 25(4): 599-634.

Strauss, J. and D. Thomas (1995) "Human Resources: Empirical Modeling of Household and Family Decisions", Jere R. Behrman and T.n. Srinivasan (Eds.), Handbook of Development Economics, Volume 3. Amsterdam: North Holland.

Thomas, D. (1990) "Intra-Household Resource Allocation : An Inferential Approach", *Journal of Human Resources*. 25(4): 635-664.

Table 1
Standard Deviations for Actual and Simulated Data by
Assumed Substitutability of Own and Co-resident's Health

	Actual Residence		Simulated Residence			
			Substitutable		Complementary	
	Within	Across	Within	Across	Within	Across
Education	2.118	1.649	1.817	1.764	1.947	1.708
Age ($\times 10^{-2}$)	0.064	0.070	0.085	0.080	0.092	0.082
Age ² ($\times 10^{-4}$)	0.079	0.066	0.088	0.078	0.090	0.083
Sons at Home	0.715	0.740	0.770	0.691	0.737	0.731
Daughters at Home	0.736	0.689	0.700	0.674	0.683	0.689
Husband's Education	1.821	1.824	1.776	1.848	1.870	1.802
Husband was Son in 1982	0.351	0.299	0.329	0.316	0.323	0.318
Current Married	0.349	0.266	0.327	0.298	0.325	0.308
Husband Absent	0.338	0.291	0.334	0.316	0.344	0.314
Husband's Age ($\times 10^{-2}$)	0.092	0.072	0.079	0.076	0.080	0.076
Husband's Age ² ($\times 10^{-4}$)	0.096	0.078	0.085	0.081	0.086	0.081
Index			0.296	0.247	0.270	0.277

Table 2
 Estimates of Household Division Model by
 Substitutability of Own and Co-resident's Health

	Substitutable	Complementary
Education	-0.083 (20.57)	-0.095 (18.89) ^a
Age (x10 ⁻²)	0.661 (0.80)	-4.059 (4.30)
Age ² (x10 ⁻⁴)	-1.660 (2.12)	4.174 (5.17)
Sons at Home	-0.137 (9.87)	-0.120 (6.32)
Daughters at Home	-0.197 (14.12)	-0.186 (7.95)
Husband's Education	-0.121 (17.34)	0.041 (4.81)
Husband was Son in 1982	-0.165 (3.88)	-0.299 (8.29)
Current Married	-0.181 (5.78)	-0.424 (5.69)
Husband Absent	-0.426 (7.77)	-0.092 (1.28)
Husband's Age (x10 ⁻²)	-0.995 (1.99)	-1.353 (3.15)
Husband's Age ² (x10 ⁻⁴)	0.630 (1.19)	1.565 (4.16)
κ_{111}	-1	-1
κ_{123}	-0.516 (16.85)	-1.530 (39.18)
Criterion Function ^b	.182	.148

^aAbsolute t-ratios based on heteroskedasticity-consistent standard errors in parentheses.

^bMean squared error computed as the product of two errors from independent simulations.

Table 3
Effects of Family Unit Characteristics on Maternal BMI Equations by Residence Status With
and Without Correction for the Selectivity of the Joint Residence Decision
Cross-sectional estimates

	Within Households		Across Households in Same Bari	
	Uncorrected	Corrected	Uncorrected	Corrected
Education	.022 (.41) ^a	-.216 (.62)	.075 (1.78)	.026 (.56)
Age (x10 ⁻²)	3.61 (.71)	3.915 (.77)	.021 (0)	5.573 (.88)
Age ² (x10 ⁻⁴)	-7.742 (1.58)	-10.986 (1.61)	-3.442 (.56)	-7.056 (1.12)
Sons at Home	-.053 (.42)	-.427 (.76)	-.063 (.77)	.13 (1.16)
Daughters at Home	-.37 (2.69)	-.858 (1.19)	.012 (.14)	.218 (1.83)
Husband's Education	.022 (.48)	-.296 (.64)	.023 (.69)	.195 (2.55)
Husband was Son in 1982	-.033 (.08)	-.252 (.44)	-.456 (2.52)	-.255 (1.81)
Current Married	.277 (.45)	-.365 (.33)	.899 (1.81)	2.066 (3.03)
Husband Absent	-.464 (.91)	-1.789 (.9)	.087 (.23)	1.378 (2.14)
Husband's Age (x10 ⁻²)	12.649 (1.91)	13.42 (1.99)	12.092 (1.92)	11.697 (1.86)
Husband's Age ² (x10 ⁻⁴)	-14.51 (2.32)	-15.637 (2.41)	-12.654 (2.2)	-14.733 (2.54)
Selection term		2.968 (.69)		2.421 (2.49)

^aAbsolute t-value in parentheses.

Table 4
Effects of Family Unit Characteristics on Maternal BMI by Residence Status With and
Without Correction for the Selectivity of the Joint Residence Decision
Longitudinal Estimates

	Intact Households		Split Households	
	Uncorrected	Corrected	Uncorrected	Corrected
Individual x				
Husband son	.134	.065	-.028	-.062
in 1982	(1.79) ^a	(0.69)	(1.54)	(2.56)
Education	.005	.004	.007	.0071
	(2.22)	(1.84)	(2.91)	(2.82)
Individual minus household average				
Husband son	.011	.056	-.0003	.022
in 1982	(0.36)	(1.13)	(0.01)	(0.76)
Education	.0005	.001	-.0002	-.000
		(0.86)	(0.34)	(0.33)
Selection term		.323		-.215
		(1.13)		(2.15)

^aAbsolute t-value in parentheses.