# Growing up together: Cohort composition and child investment

Jones, Kelly M.

International Food Policy Research Institute

2014

# Growing Up Together:

# Cohort composition and child investment

February 25, 2013

In Sub-Saharan Africa, sixty percent of child deaths are preventable by investments in child health as simple as immunizations, bed nets, or water purification. This paper investigates how a household's decision regarding such investment is affected by the size and gender composition of a child's cohort. I focus on a previously overlooked type of investment: non-rival, child-specific goods (club goods). I empirically estimate the response of immunization status to cohort characteristics. I deal carefully with the problem of endogenous fertility, common in cohort studies. Most rural Senegalese households are composed of multiple nuclear families, so that a child's cohort is composed of both siblings and non-sibling children. Estimating within-household, I instrument cohort characteristics with those of the non-sibling (exogenous) portion. I find that children with larger (or more male) cohorts of vaccine-eligible age are significantly more likely to receive immunization. These findings suggest that children with larger cohorts may be better off in terms of club investments: a significant finding for child health, as many illness prevention methods are of a club good nature.

**JEL Codes:** J13, J12, O12, O15, D13

# 1. Introduction

Evidence suggests that, from the earliest stages of childhood, investments in human capital translate into more successful adults.[1] Parents' decisions regarding investments in nutrition, health, and education for their children determine the productivity of the next generation. This is particularly salient in Sub-Saharan Africa, where the cost of even a small investment can be a significant burden, and child development lags. As such, the complex processes that affect child investment decisions in poor households are of utmost interest to studies of economic development.

Since the 19th century, researchers have recognized that such investment decisions are affected by household composition, in particular, the number of children within the household.[2] The seminal work by Becker (1960) suggests that parents trade off between quantity and quality of children, implying that children in larger families get less investment.

Becker's theory assumes that investments in children are strictly private goods, so that children are in explicit competition for resources. However, many child investments can be considered "club" goods for kids. Club goods are defined as those that are non-rival, so that, unlike private goods, one person's consumption does not diminish the amount available for others. While similar to public goods, club goods are different in that they are excludable, so that they are only available or useful to those in the "club." Any non-rival good that disproportionately benefits the health of young children over other members of the household could be considered a club good for kids. This applies to many investments that are critical for early childhood health in poor countries. Some examples include improved flooring, improved cook stoves, treated bed nets, or a method for purifying drinking water. Such investments can prevent or reduce the leading causes of childhood morbidity and mortality in the developing world: worms, respiratory tract infections, malarial illness, and water-borne diseases.

Certainly, a household with an additional child, *ceteris paribus,* will have fewer resources available per child. In this sense, a child is directly affected by the size of her cohort. I use the term cohort to mean the group of children with whom she is growing up who are *within her household,* rather than the generational meaning of the term. Given the competition for resources, a larger cohort would clearly predict reduced investment in private goods per child.

[1] See Case and Paxson (2008); Black et al. (2007); Alderman et al. (2006)

[2] Engel (1895) recognized that the "cost of a child" depends on the number of other children in the household.

However, the expected result with regard to club goods is less clear. A larger cohort would reduce available resources per member, but would simultaneously increase the returns to club investments in kids.

The question of how a cohort affects investment in club goods follows directly from the research on household economies of scale. Barten's (1964) model of household economies of scale predicts that a change in family composition, holding total income constant, will change the household's consumption choices across goods. He likens these "family effects" to price effects: greater household size reduces the per person cost of public goods. This is analogous to the proposition that a larger cohort of children effectively makes club goods cheaper relative to private goods for children. Gustafsson and Kjulin (1994) show evidence that such economies of scale do exist for adult time spent on child care, based on their analysis of families in Sweden. In the context of a poor country, such potential economies may have a significant impact on investment affordability.

Thus, with regard to club goods, a child may benefit from, rather than compete with those in her cohort. If so, club goods will respond to cohort characteristics in the opposite manner of private goods, and accounting for this response may change the net effect of one's cohort. Such issues may explain the seemingly inconsistent results when researchers have sought causal evidence of the quantity-quality trade off.

## Contribution

The primary contribution of this work is to estimate, in the Sub-Saharan African context, the effects of cohort size and gender composition on investment in club goods for children. The central challenge in estimating such effects is that the parents making the investment decisions have also (to some degree) chosen these cohort characteristics. Therefore, many unobserved characteristics and preferences will affect both cohort and investments.

This endogeneity of siblings plagues many empirical studies of cohort effects. Previous efforts to deal with this endogeneity have employed twinning as an exogenous determinant of total fertility.[3] Though a creative and rigorous solution to the endogeneity problem, it is difficult to know whether results are applicable to non-twin families. Further, this methodology is infeasible in the poor country context, as twins data sets are generally unavailable. Other researchers have also employed China's one-child policy as an exogenous determinant of cohort

[3]See Black et al. (2010, 2005); Caceres-Delpiano (2006)

size.[4] For this method too, it is difficult to know whether results would generalize outside the Chinese context.

A second contribution of this paper is the introduction of a new method for dealing with this endogeneity. I exploit the large households that are the norm in Senegal, generally composed of extended families and containing multiple mothers of young children. These are generally sisters-in-law. In these households, a child's non-sibling co-resident children provide similar cohort effects with regard to club goods as would her siblings. However, when estimating within-household, these non-sibling co-resident children are exogenous with respect to the fertility decisions of the child's mother. The characteristics of this exogenous portion of the cohort are used as an instrument for the full cohort measures in a two-staged least squares estimation.

In the following section I discuss the existing literature regarding cohort effects on child investment. Section 3 presents the empirical analysis and section 4 concludes.

## 2. Existing Evidence of Cohort Effects

The theory of a quantity-quality trade-off in child rearing, first proposed by Becker (1960) and then formalized by Becker and Lewis (1973), suggests that investments in children are decreasing in the size of their cohort. It can be easily shown in a broad range of settings that family size and child outcomes exhibit a negative correlation. However, few researchers have been able to establish a *causal* effect of cohort size on investments. For example, studies of the Matlab family planning experiment in Bangladesh do examine an exogenous change in fertility, however the implementation of child health services alongside family planning services makes it impossible to ascribe the estimates of improved child health outcomes (or investments) solely to reduced family size (Joshi and Schultz, forthcoming).

The most common way to deal seriously with the endogeneity of fertility choices in this literature is to exploit the exogenous increase in fertility that occurs with twinning. However, the results presented by the "twins studies" regarding the impact of cohort size on child outcomes has been mixed. While Rosenzweig and Wolpin (1980) and Rosenzweig and Zhang (2009) find decreased schooling for children from families with twins, Angrist et al. (2010); Black et al. (2005, 2010); Caceres-Delpiano (2006); Qian (2009) do not find significantly neg-

[4] See Rosenzweig and Zhang (2006); Qian (2009)

ative effects of increased cohort size on schooling or health outcomes. Employing an alternative method, Rosenzweig and Schultz (1987) estimate a woman's fecundity and use it as an instrument for total fertility in Malaysia, where fertility control is incomplete. This study shows that increased cohort size reduces schooling.

These studies generally focus on outcomes for children, rather than investments *per se*. Outcomes such as schooling attainment or achievement are, of course, the result of both private and club investments. As such, if cohort effects that decrease investment in private goods are offset by increased club goods, one would expect these ambiguous or indistinguishable effects on such aggregate outcomes. One exception is a secondary result in Rosenzweig and Zhang (2009), where the authors examine actual investments such as adult time spent helping with homework and the likelihood of having an internet connection. One could consider both of these investments club goods, and in both cases the authors find a positive effect of cohort size.

In addition to cohort size, a child's own gender may also affect parental decisions regarding allocation of resources. Several empirical studies have shown evidence of son preference in child investment in diverse cultural contexts.[5] In the presence of son preference, boy-children may command a greater share of household resources than girl-children, leading some scholars to suggest that girls who have more brothers are worse off (holding total cohort size constant). That is, assuming that parents allocate a fixed share of resources to children, the resources available for a given child are decreasing in the share of her cohort that is male.

Evidence for such gender "crowding out" has been mixed. While some studies that focus solely on education have found negative effects of a male cohort,[6] effects on other composite outcomes such as health and mortality are ambiguous. In India, Makepeace and Pal (2008) show evidence of gender crowding out, in that boys with more male siblings have higher mortality rates. But both Pande (2003) and Mishra et al. (2004) examine anthropometric measures for children in India and find that in some cases girls are actually disadvantaged by sisters, and not brothers.[7] In Ghana, Garg and Morduch (1998) find that anthropometric measures for

[5]See studies based in Nepal, Japan, Egypt, the U.S., China, and India: Edmonds (2006); Ono (2004); Yount (2003); Lundberg (2005); Gong et al. (2005); Asfaw et al. (2010); Duraisamy and Duraisamy (1995); Rose (2000), respectively.

[6]See Parish and Willis (1993) in mid-century Taiwan; Morduch (2000) in Tanzania & South Africa; Bommier and Lambert (2004) in Brazil; Ota and Moffatt (2007) and Kambhampati and Rajan (2008) in India; Rammohan and Dancer (2008) in Egypt; and Dayioglu et al. (2009) in Turkey.

[7]The authors' explanation is that selective neglect arises from a desire for gender balance,

children are 25-40% worse for a child with an all-boy cohort (versus all girls). However, the authors also find a contrasting positive effect of the indicator for having *any* brothers. Such seemingly contradictory results might be explained by corresponding increases in club goods investment and decreases in private investment in response to male composition of the cohort.[8]

# 3. Empirical Analysis

The purpose of the empirical exercise is to test how investment in club goods responds to cohort size and gender composition. While some club goods for children are physical, such as a household water sanitation system, many non-rival investments in children are composed of adult time in some form (e.g. supervision, reading, transportation, etc.). The example club good employed here is the transportation of children from rural areas in Senegal to a health clinic for immunizations. While immunization may be seem to be a private good, the injection itself is generally costless in this context due to government subsidization (SaveTheChildren, 2012). However, the adult time required to obtain the immunization (via transport and wait time) may be many hours. Pande (2003) notes that, in India, differential immunization rates by gender, despite decades of free provision, imply significant opportunity costs of adult time. Further, it seems likely that parents traveling to the clinic would bring all children in the household in need of an immunization. Therefore, as the explicit cost of the injection gets close to zero, a vaccine becomes (very nearly) a non-rival good for children of a relevant age. In section 3.2, I provide suggestive evidence that immunizations are a club good in the context of this analysis.

## 3.1. Data

This analysis is set in Senegal, owing to the Senegalese households that are large and often composed of multiple families. This characteristic enables me to resolve an inherent challenge of estimation, as discussed in section 3.3, and is not unusual in the developing world. Senegal is an appropriate context for this analysis, having a total fertility rate more than double the world average (5.2 children per woman in 2005), so that children have large cohorts of varying sizes.

rather than a pure desire for sons.

[8]The authors instead attribute this unexpected result to psychological phenomena: "reference effects" for boys, that is, a lone boy with all sisters is treated more like a girl; and "spillover effects" for girls, that is, girls adopt and/or are taught more "masculine" traits in the presence of a brother.

There is also evidence of preferences for having and investing in sons relative to daughters in this context. Household compositions suggest the presence of differential stopping behavior, a practice induced by a preference to have sons (discussed further in section 3.3). This is supported by the self-reported fertility preferences, where average ideal number of boys is significantly higher than ideal number of girls (2.9 vs. 2.2; *p-value* = 0). Further, stronger investment in boys is evidenced by the higher rate of immunization for boys in these data (by 0.5 ppt, *p-value* = 0.03).

The data are drawn from two cross-sectional household surveys, the Senegalese Demographic and Health Surveys (DHS) in 1993 and 2005.[9] The DHS interviews all women aged 15-49 from a nationally representative sample of households on topics relevant to fertility, reproductive health, marital relations, and childhood health and nutrition. Specifically, the two surveys employed here recorded whether each child under age five did or did not receive each vaccine recommended by the Senegalese National Immunization Schedule as shown in Table 1. While it is possible that a single household appears in both data sets, it is not possible to link households across data sets. Therefore, the two cross-sections are pooled for estimation purposes. However, given that the estimation is based only on children under five, no child appears in both sets of data, so that each child-observation is truly unique.

Table 1 also presents a summary of the coverage of immunizations in rural households. Overall coverage of immunizations is low, though increasing from 60% to 75% over the 12 year period. Vaccines for which a child is not yet due are excluded from the mean, giving rise to the slight decline in sample size for vaccines due at older ages. Further details regarding sample size are show in Table 2. A decline in sample size results also from non-response to some items regarding immunization. However, children with data missing for any of the ten vaccines make up less than 10% of the sample. The last two rows of Table 2 show the sizes for the two samples used in most specifications: 92,643 child-vaccine observations, and 50,602 child-vaccine observations for children with a non-zero cohort size.

[9]Other SDHS rounds were not used for the following reasons: 2008, 2006 and 1997 did not collect immunization information; 1986 collected immunization information only for a small sub-sample.

## 3.2. Immunizations as a club good

Why would immunizations act as a club good in this setting? I argue here that the principal cost of immunizing a child is the cost of adult time to wait at the clinic, which, in addition to travel time, is non-rival for any children who need to be taken. Certainly, immunizations are not perfectly non-rival, as there may be some positive marginal cost per child, however these are likely small relative to the total cost. One expects that mothers' transportation to the clinic is primarily by foot (see discussion below), so that walking children will exert zero marginal cost. Non-walking children will exert some positive marginal cost in terms of adult effort, though nearly zero in terms of adult time. However, there is an upper bound of non-walking children that can be carried, so we expect that the cost-per-child associated with immunizations as a club good would be decreasing in cohort size, but at a decreasing rate. This is supported by results presented in section 3.5.2 that suggest a non-linear relationship between immunizations and cohort size.

So, what would be the total cost in terms of adult time? Community-level data collected with the 1993 DHS include information on distance to the nearest health facility. On average, these rural communities are 5.4 km from a service access point. At this distance, most individuals would travel by foot, equating to a mean travel time of 81 minutes, one-way.[10]

Systematic data on clinic wait times in Senegal are not available. However, one document reports average visit time for each of four urban family planning clinics in Senegal, and these averaged nearly three hours (171 minutes) (SEATS, 2000). This is consistent with evidence from other African countries: client wait times in two clinics in Nairobi, Kenya were 140 minutes, for only 16 minutes of service time, on average (Lynam et al., 1994); wait time at an urban HIV treatment clinic in Nigeria was 6.48 hours on average (Umar et al., 2010). Given the lower provider-to-client ratio in rural areas, an appointment at a rural clinic could easily take half the day or more.[11] In a study of rural health clinics in Mozambique, 43% of patients reported that either the lengthy travel or the long wait time was the most significant barrier affecting their utilization (Newman et al., 1998). Based on the evidence presented here, it seems reasonable to expect that a rural Senegalese mother would require five to seven hours for

[10] See Tanser et al. (2006), which uses GIS and transportation modeling for rural Sub-Saharan Africa and predicts that for distances of 5 km, only 20% of individuals will use public transportation.

[11] See Clemens (2009), pages 15 and 16.

accessing vaccination services, a large cost for a service with an unseen benefit.

In this analysis, immunizations are considered a club good for children in a relevant age group. Although the vaccines considered here are all due by age 9 months, in practice, the children in these data continue receiving vaccines up to 4 years of age. However, over 95% of vaccines administered are received by age 2 years, and this will be the default age group used in this analysis.[12] Therefore, for a given child, *i*, the relevant **age-cohort** is composed of all other children in the household within 24 months of his age; that is, co-resident children potentially needing vaccines at any time during which child *i* is due for a vaccine.[13]

As evidence that immunizations operate as a club good, I examine the proportion of children receiving vaccines on the same day as a cohort member. For approximately half of the full sample, the enumerator was able to view the health card that lists the immunization history of a child. For this health card sub-sample, the data contain the date on which each immunization was received. Of rural children with age-mates, 44% of those that have completed their vaccines received one or more at the same time as another child in the household. In contrast, urban children are 30% less likely to do so ($p - value = 0.00$). This difference likely reflects the increased opportunity cost involved for rural residents to reach and wait at a health clinic. This motivates the restriction of this analysis to the rural areas and suggests that, in rural areas, travel to the clinic is a non-rival good for children in need of immunizations.

## 3.3. Methodological Challenge

In any empirical estimation of the response of allocation to composition, one must account for the fact that neither the number of children in the household nor their age distribution is exogenous. Specifically, unobserved factors that determined parents' past fertility choices will also impact current allocation decisions.

It is uncontroversial to state that parents choose, to some degree, the number of children they bear. Perhaps less obvious, but also true, is the statement that parents choose the gender composition of their children as well, within some bounds. If parents have greater desire for sons (for any reason), it is likely that childbearing will continue until some target number of sons is achieved. Yamaguchi (1989) presents a formal model of how son preference can

[12]Results are robust to age groups defined over the range of 12 to 36 months.

[13]This terminology is not meant to imply any relation to age-period-cohort analysis.

affect the gender ratio even when manifest only through such differential stopping behavior (DSB). DSB predicts that girls will end up with more siblings than boys, and that families with greater son preference will have a more male-skewed composition. Filmer et al. (2009) present empirical evidence that DSB is practiced on a significant scale, based on DHS data from 65 low-income countries. While DSB seems most prevalent in South and Central Asia, evidence suggests that DSB is practiced in rural Senegal. In the data employed here, boys are significantly more likely to be last-born than girls, suggesting that the birth of a girl is more likely to induce a continuation in childbearing. Girls in this sample have more sibling age-mates on average than boys, and a negative correlation between parity and male-composition is observed.[14] When DSB is practiced to any degree, it creates correlation between parental preferences and observed gender composition of children.

Such correlations between preferences and both number and gender of children can seriously confound efforts to estimate how child-composition affects intra-household allocations or child outcomes. Ejrnaes and Portner (2004) simulate the relationship between birth order and schooling investment in children based on a model of household allocation in which fertility is endogenous. Their results show that accounting for endogenous fertility reverses the direction of correlation between birth order and schooling; that is, when fertility is assumed exogenous, empirical estimates are seriously biased.

*Solutions*

In order to avoid confounding by unobserved household characteristics, I employ a within-estimator using household fixed effects to compare across children within the same household. In combination with an indicator for a child's birth order, this holds constant the total number of children in the household when the child is due for immunizations, as well as time-invariant household characteristics. Note that this is possible only because the cohort of interest is defined by age and therefore varies across children within the same household.

However, since many households in the sample include multiple mothers of young children, one may still be concerned that unobservable differences between mothers within a house-

[14] Girls have an average of 0.49 sibling age-mates vs. 0.46 for boys; equality rejected with *p-value* = .0539. Among women with reasonably completed fertility (age 48+), each additional child is associated with a reduction in male share of .022 (on a mean of .47, *p-value* = 0.03).

hold could confound the estimation.[15] For example, a mother that has a high concern for child health may be careful to space births according to public health recommendations (at least 3 years apart). She would also be sure to immunize her children, creating a spurious negative correlation between age-cohort size and immunizations. Mother-fixed effects could remedy this problem but are unfortunately not feasible in this sample. Variation in age-cohort composition across a mother's children would require her to have at least 3 children in the sample. Since vaccination information is only available for children born within 5 years of the survey, fewer than 10% of mothers have 3 or more children in the sample. This is, of course, merely a problem of lacking data that contain detailed immunization information for *all* of a woman's children. Further data collection would allow a more ideal estimation. What is presented here is a novel method for achieving causal estimates based on data that are currently available.

Mother-fixed effects being infeasible, I contend with this potential endogeneity by exploiting a unique characteristic of households in Senegal. As shown in Table 3, 56% of mothers of children under age five co-reside with one or more other mothers of young children.

*Multi-family households* Who are these co-resident families? Explicit relations between household members are not given in the data; only each's relation to the household head is given. However, for households that have more than one mother of young children, I can partially determine which are resulting from polygyny according to the method described in the notes of Table 4. While some of these households result from polygynous unions, three-quarters of them do not. Compared to other rural women, women in households with multiple mothers (that are not co-wives) are much more likely to be the daughter-in-law or a foster daughter of the household head, as shown in Table 4. This suggests that many co-resident mothers are often sisters-in-law, foster-sisters, or other relatives.

I assume that a family is unitary, so that what I describe as mothers' decisions throughout the text are actually the outcome of some joint decision-making or bargaining process between mother and father. I also assume that within a household, separate families coordinate on the provision of goods and services for the household, but do not coordinate on matters of sex and procreation.[16] Evidence that co-resident families do not coordinate fertility timing is provided

[15] In this analysis (as in the data), a family is defined as a mother and her children. A household is defined as members who co-reside – generally extended families, as discussed in the following sub-section.

[16] Of course, the latter assumption is in conflict in the case of polygyny. Implications of this are

in appendix A.

If families coordinate on provision of goods, a trip to the health clinic would be non-rival for all children of a eligible age within the household. Therefore, the benefit of time spent traveling to and waiting at the clinic will have a value that depends on the total number of eligible children in household. Thus, all age-mates in the household (sibling or not) would be the relevant cohort for evaluating the cohort composition effects of club good consumption.

As discussed above, the sibling portion of the cohort is endogenous. However, by employing household fixed effects, I hold constant any unobserved preferences held in common within the household, so that the timing of non-sibling co-resident children is exogenous. I can then use the number and gender of the non-sibling age-mates as an instrument for the characteristics of the full cohort.

*The primary identification assumption is thus:* controlling for anything common across mothers in a household by the use of household-fixed effects, the immunization decisions of mother *A* are impacted by the fertility decisions of mother *B only* through the economies of scale effects that are proposed here. If any unobserved preferences that affect both fertility and immunization are shared across mothers (because they share a husband, for example), this will be absorbed by the household fixed effect.

Three possible threats to this identification exist. The first is the case where there are three or more co-resident mothers in a household and at least two, but not all, of them share a husband. In this case the two wives sharing a husband could have correlated preferences that are not captured by the household fixed effect. However, as shown in section 3.5.3, the results are robust to the exclusion of households potentially fitting this unique profile.

The second case is similar, which is the possibility that the assumptions about household dynamics are wrong. What if, rather than unitary families that are independent within households, the households themselves are in fact unitary. If there is a single patriarch making the primary decisions for everyone in the household, then these preferences or decisions will be captured by the household fixed effect. In this case, there are certainly some, smaller decisions that are taken independently by mother-father pairs, and it is these smaller decisions that will be the source of variation and identification for my estimations. In appendix B I show some basic evidence that there is heterogeneity within households regarding immunization decisions, indicating that there is some parental autonomy within the household.

discussed in detail below and in section 3.4.

The third is a threat of reverse causality – perhaps the immunization decisions of mother $A$ are affecting the fertility of mother $B$. That is, if $A$ has a strong preference for investing in club goods for children and will certainly do so, $B$ may calculate that her own costs of investing in her children are lower (due to free-riding). If these savings reduce the expected costs of child-rearing significantly, it is possible that this may induce mother $B$ to coordinate her fertility timing with mother $A$. In this case, a similar positive correlation between mother $A$'s investment and mother $B$'s co-timed fertility would be observed, but the direction of causation would be opposite. However, as shown in appendix A, conception is not significantly predicted by recent conception of a co-resident woman in non-polygynous households, suggesting that co-resident women are not coordinating fertility timing, either to have co-incident or spaced births.

## 3.4. Estimation

Absent the endogeneity issues discussed above, the model I would like to estimate is

$$R_{ijv} = \alpha + \beta_1 N_{ij} + \beta_2 Z_{ij} + \beta_3 K_{ij} + \eta_j + \eta_v + \varepsilon_{ijv} \quad (1)$$

where $R_{ijv}$ is a binary indicator of whether child $i$ in household $j$ received vaccine $v$. The outcome is predicted by $N_{ij}$, the size of child $i$'s age-cohort within the household, and $Z_{ij}$, the share of that cohort that is male, child characteristics contained in the vector $K_{ij}$ as discussed below, and household and vaccine fixed effects, $\eta_j$ and $\eta_v$. $\varepsilon_{ijv}$ is a mean-zero error term. However, given the potential endogeneity of $N_{ij}$ and $Z_{ij}$, the model I estimate instead is

$$R_{ijv} = \tilde{\alpha} + \tilde{\beta}_1 \tilde{N}_{ij} + \tilde{\beta}_2 \tilde{Z}_{ij} + \tilde{\beta}_3 K_{ij} + \eta_j + \eta_v + \tilde{\varepsilon}_{ijv} \quad (2)$$

where $\tilde{N}_{ij}$ and $\tilde{Z}_{ij}$ are the analogs to measures $N_{ij}$ and $Z_{ij}$ for the exogenous portions of the age-cohort (the non-sibling co-resident children). For comparison purposes, I also estimate the same equation employing only the endogenous portion of the age-cohort (the age-mate siblings). That is,

$$R_{ijv} = \dot{\alpha} + \dot{\beta}_1 \dot{N}_{ij} + \dot{\beta}_2 \dot{Z}_{ij} + \dot{\beta}_3 K_{ij} + \eta_j + \eta_v + \dot{\varepsilon}_{ijv} \quad (3)$$

where $\dot{N}_{ij}$ and $\dot{Z}_{ij}$ are the relevant size and gender composition measures for the sibling age-cohort. In this equation, $\tilde{\beta}_1$ ($\dot{\beta}_1$) represents the predicted change in the probability of receiving

a vaccine for each additional non-sibling (sibling) member in a child's age-cohort. The coefficient $\tilde{\beta}_2$ ($\dot{\beta}_2$) represents the difference in that same probability for a child who has non-sibling (sibling) age-mates that are all boys versus all girls. In this setting, I assume that the true effects are the same for sibling and non-sibling age-mates within the household, that is, $\tilde{\beta}_1 = \dot{\beta}_1$ and $\tilde{\beta}_2 = \dot{\beta}_2$. However, I expect that the coefficient *estimates* based on sibling age-mates will be confounded by unobserved maternal characteristics, so that $\tilde{\beta}_1 = \widehat{\tilde{\beta}}_1 \neq \widehat{\dot{\beta}}_1$ and $\tilde{\beta}_2 = \widehat{\tilde{\beta}}_2 \neq \widehat{\dot{\beta}}_2$. That is, I expect that estimates of $\dot{\beta}$ will be biased, but that the estimates of $\tilde{\beta}$ will be unbiased. If parental concern for child health induces both lower fertility and increased investment, the bias on $\widehat{\dot{\beta}}$ will be downward.

While OLS estimation is useful for recovering unbiased estimates of $\tilde{\beta}$ based on the non-sibling age-mates, I am rather more interested in estimates of $\beta$, that is, the effects of the characteristics of the *full* age-cohort. Thus, I subsequently employ a two-stage least squares estimation to recover the $\beta$ coefficients. I use the exogenous portion of the age-cohort to predict the measures $N_{ij}$ and $Z_{ij}$ for the full age-cohort. I estimate

$$\hat{Z}_{ij} = \delta_0 + \delta_1 \tilde{Z}_{ij} + \delta_2 \tilde{N}_{ij} + \delta_3 K_{ij} + \eta_j + \eta_v + u_{ijv} \quad (4)$$

$$\hat{N}_{ij} = \theta_0 + \theta_1 \tilde{Z}_{ij} + \theta_2 \tilde{N}_{ij} + \theta_3 K_{ij} + \eta_j + \eta_v + v_{ijv} \quad (5)$$

$$R_{ijv} = \alpha + \beta_1 \hat{Z}_{ij} + \beta_2 \hat{N}_{ij} + \beta_3 K_{ij} + \eta_j + \eta_v + \varepsilon_{ijv} \quad (6)$$

so that $\widehat{\beta}_1$ and $\widehat{\beta}_2$ have similar identification as $\widehat{\tilde{\beta}}_1$ and $\widehat{\tilde{\beta}}_2$, but are now appropriately scaled to represent effects of the full age-cohort.

Given that an increase in cohort size would reduce available resources per child but also reduce the price of club goods per child relative to private goods, the expected sign of $\widehat{\beta}_1$ will be the net of income and substitution effects. $\widehat{\beta}_1$ will be positive if the substitution effect dominates the income effect. Similarly, the expected sign of $\widehat{\beta}_2$ is ambiguous. An increase in the male share of the cohort makes club goods relatively cheaper than private goods to boys, but more expensive than private goods to girls. Absent any son preference, this would predict that resources would be diverted from boy goods to girl goods ($\widehat{\beta}_2 = 0$), or possibly also diverted from club goods to girl goods ($\widehat{\beta}_2 < 0$). However, households with son preference would prefer to shift resources from boy goods to club goods that also benefit boys, rather than goods benefiting only girls. $\widehat{\beta}_2$ will be positive if son preference dominates the price effect.

A final concern for the estimation of $\beta_2$ is that $1/3$ of the children in the sample have no age-mates within the household. For these children, the gender composition of their age-cohort is meaningless. Therefore, in order to accurately estimate $\beta_2$, the effect of gender composition, I restrict the sample to only those children with a non-zero cohort size (NZC sample). However, to accurately measure $\beta_1$, I employ the full rural sample and thus must exclude $Z_{ij}$. In result, for each of the above equations, two estimations are shown: the full sample, excluding $Z_{ij}$, and the restricted NZC sample, including both $N_{ij}$ and $Z_{ij}$.[17]

*Control variables* Certainly there are some child characteristics that effect a child's probability of immunization beyond her cohort size and gender composition. To the extent that any of these child characteristics are also correlated with these cohort characteristics, I am careful to include them so that their impact does not confound the estimate of interest. The primary characteristic of concern is the child's gender, as boys are more likely to be immunized and, due to differential stopping behavior, boys also have slightly smaller and more male-dominated age-cohorts than girls (see the opening of section 3.3 for further discussion). I thus control for child gender to ensure that the impact of age-cohort size is not under-estimated and that the impact of age-cohort male composition is not over-estimated. Both birth order to a mother and birth order within the household also predict both immunization and age-cohort size, to lesser degrees. Children later in the birth order to a mother or within the household are less likely to be immunized. Children later in a mother's birth order have smaller age-cohorts. Children later in the household birth order have larger age-cohorts. As such, these controls are also included in the fully specified model. Finally, while age is also a predictor of immunizations, I do not include it due to a high degree of collinearity with mother's and household birth order. It is shown in table 8 that the main results are not dependent on the inclusion of of these controls.

## 3.5. Results

Table 5 shows results from estimations of eq.s 3, 2, and 6. Columns 1, 3, and 5 employ the full rural sample in order to estimate the effect of cohort size. Columns 2, 4, and 6 employ the sub-sample that has a non-zero cohort size (NZC sample) in order to recover the effect of

[17]Note that the exclusion of $Z_{ij}$ from the full sample estimation may slightly downward-bias the estimate of $\beta_1$ , as $N_{ij}$ and $Z_{ij}$ are correlated due to differential stopping behavior (see subsection 3.3 and footnote 14).

cohort gender composition. Columns 1 and 2 show estimates of eq. 3, taking only a child's siblings as her cohort. As noted above, it is expected that such estimates will be significantly downward biased. These show negative point estimates for the effect of cohort size and a small positive effect of male composition that is not distinguishable from zero.

The results in columns 3 and 4 employ only the exogenous portion of a child's cohort – the non-sibling age-mates. While these children would exert similar cohort effects within the household, their presence is uncorrelated with unobserved parental preferences (when employing household fixed effects). The estimated effects are significantly different from the biased estimations employing siblings. Each additional non-sibling age-mate predicts an increase in probability of immunization of about 2.5 percentage points. This estimate is robust to whether or not I control for the gender composition, and whether or not children without any age-mates are excluded. In column 4, the sample includes only children that have exogenous age-mates in order to examine gender composition effects. It suggests that, controlling for cohort size, having an all-male cohort predicts a 3.6 pp greater probability of immunization than having all female age-mates.

To estimate the effect of an age-mate more generally, columns 5 and 6 show estimates that instrument full cohort characteristics with those of just the exogenous portion (non-sibling age-mates). These are the second-stage estimations of eq. 6.[18] The coefficients of interest have increased in magnitude relative to the OLS estimation, though the 95% confidence intervals do overlap. These results suggest that any additional age-mate will increase the probability of immunization by about 3pp. Further, for children having any age-mates, an all-boy cohort predicts a 4.6pp increase in the likelihood of immunization relative to an all-girl cohort.

Given an immunization coverage rate of only 70%, the magnitudes of these effects are significant. Consider a child that has one female age-mate with $R = 0.70$, where $R$ is the probability of receiving a given vaccine. Then a child with four female age-mates would have $R = 0.79$, an increase of 13%. Further, if those age-mates were all boys, then $R = 0.836$. That is, she would be nearly 20% more likely to be immunized than the child in the baseline case.

### 3.5.1. *Heterogeneous effects*

Table 6 repeats the 2SLS estimations for various samples that are assumed to have different costs associated with taking children for immunizations. The greater the share of the cost that

[18]Estimations of first-stage eq.s 4 and 5 are shown in Appendix C.

is non-rival, the more immunizations are a club good. For example, having access to a health clinic within one's village significantly reduces the travel time required to receive immunizations, thus reducing the share of the cost that is non-rival. Thus, households with a clinic in the same village should exhibit smaller effects of cohort size. Unfortunately, distance to clinic information is only available in the 1993 round of data, but nonetheless this heterogeneity be examined within the 1993 sample.[19]

Columns 1 and 2 of Table 6 show estimates for households in villages that reported having a health clinic zero kilometers away. For both estimates employing the full sample and the sub-sample of children with a non-zero cohort size (NZC sample), the impact of cohort size is reduced by about 50% and is no longer statistically distinguishable from zero. Within the NZC sample, the impact of cohort gender composition is inflated. It is not clear why households nearer to clinics would exhibit a stronger impact of gender composition, except for the possibility of differential son preference. However, the point estimate is also less precise and not distinguishable from zero at any standard level of significance.

Another factor that determines the non-rival share of cost for immunization is the opportunity cost of adult time. For mothers with a higher cost of time, the share of immunization cost that is adult time increases. I separate children by an indicator of the potential opportunity cost of adult time: mother's level of education. Unfortunately, a vast majority of mothers report no education, so the sample for those with any education is small. Nonetheless, we find that the point estimate on cohort size more than doubles for those whose mother has any education versus uneducated mothers, remaining statistically significant at the 1% level. The coefficient on gender composition also increases, by 70%, though has an increased standard error and I can no longer reject that it is zero. Though the standard errors do not allow us to reject that the effect is the same across the two groups, the increased point estimates are consistent with a higher valuation of time for women with education (based on a higher potential for gainful employment), and thus a greater share of the immunization cost that is non-rival.

In sum, I interpret this as evidence that the impacts of cohort characteristics on immunization are rooted in the club good nature of immunizations and their potential economies of scale in terms of adult time.

[19]Main results are disaggregated by survey round are shown in appendix D. These are consistent with the aggregated estimates, however the impact of gender composition is not statistically significant when using the 2005 data alone.

### 3.5.2. *Specification checks*

Table 7 presents several alternative specifications of eq. 6, to check for significant interactions or non-linearities. Column 1 suggests that the effect of gender is linear with respect to the share of the cohort that is male. Neither the indicator nor its square are statistically significant, though they are jointly significant. There is weak evidence, however, of diminishing marginal returns to cohort size. It seems sensible that there would be a limit on the number of children one adult could take to the clinic at once, which would produce this shape. The point estimates suggest a concave response function that levels off around 10, turning around completely at 12.25. While diminishing marginal returns seem plausible, one would expect the function to level off somewhat earlier. However, the squared term is not very precisely estimated and a turn-around at 3.7 is within the 90% confidence interval.

In column 2, the specification includes interactions of the variables of interest with the gender of child *i*. The signs of the coefficients on the interaction terms suggest that cohort effects may be slightly stronger for girls than boys. However, these are not at all estimated precisely and thus we cannot reject that the effects are the same for boys and girls.

Column 3 investigates whether there is any interaction between the effects of cohort size and cohort gender composition. The positive coefficient on the interaction suggests (i) that the effect of an additional age-mate is larger if that age-mate is a boy, and (ii) the effect of gender composition is stronger for larger cohorts. Yet again, we cannot reject that there is no interaction whatsoever, given the rather large standard error on the interaction term. Note that the lack of significance of the coefficient on male share is inconsequential, as that estimate represents the effect of cohort gender when the cohort is of size zero. That is rather meaningless; it is more meaningful to note that the coefficients on male share and the interaction are jointly significant at the 5% level.

### 3.5.3. *Robustness Checks*

*Sample selection on survival* The DHS data employed here contain immunization for all *living* children under age five. Therefore, children born within the past five years that have since died are present in the data but have no immunization information and are thus excluded from the estimations. If we assume that the excluded children received immunizations with the same likelihood as surviving children, this sample selection should not be a source of bias.

However, immunizations are intended to prevent potentially fatal illnesses. Therefore, a

negative correlation between vaccine status and child death is possible. Further, if a child is born weak or sickly, she may be less likely to be immunized, and also less likely to survive. Yet in order to bias the results presented here, death must also be correlated with size and/or gender of one's age-cohort. If the excluded children had significantly smaller (or more female) age-cohorts, then these results are downward bias and serve as a lower bound effect. If excluded children had larger (or more male) age-cohorts, than these effects are overestimates.

As shown in col. 1 of table 8, I find that deceased children had smaller age-cohorts, suggesting that the results presented here are a lower-bound on the effect of cohort size on immunization. There is also a negative, though insignificant correlation between survival and gender composition of one's cohort.

*Polygyny* The end of section 3.3 presents a possible violation of the identification assumption arising from polygynous marriages where mothers of young children share a husband. Note however that if all the mothers of young children in the household share the same husband, then the commonalities in preferences are captured by the household fixed effect. The potential violation occurs when there are (at least) two mothers that share a husband and (at least) one other mother with a different husband. There are 119 households in the sample that meet these criteria. The col. 2 of Table 8 shows the estimation of eq. 6 based on children with exogenous age-mates, excluding the households that potentially violate the identification assumption. The coefficient estimates of interest are not significantly changed by this exclusion.

*Within-household health externalities* Many of the vaccines examined here prevent diseases that are transmitted human-to-human, either via airborne transmission (Tuberculosis, Pertussis, Measles, Diphtheria), direct contact (Diphtheria), or the fecal-oral route (Polio). For these, one could imagine that children with larger cohorts are simply at higher risk of disease, given their increased potential exposure.[20] If so, then perhaps these results can be attributed to this unique nature of vaccines, rather than to the club nature of the investment.

I test for whether this could be driving the results by exploiting data for one vaccine that protects against vector-borne, rather than human-to-human transmission.[21] Yellow fever is con-

[20]There is also evidence that diseases transmitted within the household can be more severe than the same disease contracted from a casual contact (Aaby et al., 1984).

[21]Tetanus is also not passed by human-to-human transmission, but unfortunately the vaccine for Tetanus is combined with Diphtheria and Pertussis in the DPT series.

tracted from the bite of an infected mosquito. Certainly, proximity to other infected humans is a factor in vector-borne transmission. However, *Anopheles gambiae*, the primary malaria vector in Africa, has a maximal flight distance of 10km (Kaufmann and Briegel, 2004). This implies that exposure is based on infection rates within a village more so than within the household, so that vaccinating against yellow fever should offer less within-household health externalities. If such externalities are driving the results, I should find a lower point estimate for cohort size when estimating based on yellow fever immunization alone, as compared to the main estimations for all vaccines. As shown in column 3 of Table 8, the point estimate for cohort size is identical to that from the main estimation, with a very similar standard error, retaining statistical significance at the 5% level.[22] Neither is the estimate for gender composition statistically distinguishable from the originally estimated effect, though it is less precise due to the 90% reduction in sample size for this estimation.

Given that within-household transmission can result in increased disease severity, regardless of transmission mechanism (Aaby et al., 1984), I cannot fully rule out that some externalities may be at work. Nonetheless, if such externalities were fully driving the results we would expect the results to be stronger for diseases transmitted by human-to-human contact, and we do not find that this is the case.

*Age-cohort definition* Another assumption of this empirical test has been that the age-cohort that is relevant for immunizations as a club good is defined as all children in the household within 24 months of one's age. As shown in the remaining columns of Table 8, whether one defines age-mates as children within 36, 30, 18, or 12 months of one's age (columns 4 - 7), the results do not differ significantly from those based on the 24 month definition. Coefficients on cohort gender range from .035 to .062; none are significantly different from the originally estimated .046 (or from each other), and all are different from zero at standard levels of significance. Similarly, effects of cohort size range from .014 to .038, not differing significantly from the originally estimated .029. Though only the estimations based on the NZC sample are shown (in order to show the gender effects), estimates of cohort size effects based on the full

[22] One would not expect this pathway to generate the main results found regarding cohort gender composition. Nonetheless, I include the full specification, including gender composition, in the presented results based on yellow fever alone. The point estimate for gender composition (.05) is similar to the original (.036) and robustness-checked (.044 and .063) estimates, though the standard error is slightly higher ($p - value = .102$).

sample generally do not differ from the estimates shown in columns 4 - 7. The only exception is the precision of the coefficient estimate for cohort size under the 12 month definition. The final column of Table 8 shows that the effect of cohort size is statistically different from zero in the full sample; the lack of precision in the 12mo-NZC sample is likely due to the reduced sample size.

## 4. Conclusions

In this study I have considered how investment in children is affected by the number and gender composition of children within a household. Previous literature has offered (at least) two theories on this. One, that children receive less investment as their cohort increases in size; that is, Becker's theory of quantity-quality trade-offs. And two, given any preference for sons, children with more boys in their cohort will receive relatively less; that is, gender crowding-out.

Evidence for these theories based on rigorous empirical work has been mixed, at best, tending to reject more often than support them. I've proposed that, in focusing exclusively on competition for private goods, these theories are missing a key element of child investment. That is, investment in children comes in the form of both private and club goods for children. Further, the effects of one's cohort on goods provision can work in opposite directions for these two types of goods. Due to economies of scale, investment in club goods may increase as their price relative to private goods decreases under larger cohorts.

It is important to note that many significant investments in children are not strictly private in nature. The primary causes of childhood morbidity and mortality are diarrhea, acute respiratory infections, malarial illness, and parasitic worms. Such ailments result from the lack of key investments that are non-rival and disproportionately benefit the children of the household: immunizations, water purification, improved cook stoves, treated bed nets, and improved flooring. As such, it is crucial to understand how households make decisions regarding club good investments, especially as they may differ significantly from private goods.

I've shown here evidence that investment in a certain club good, immunization, is increasing in cohort size and male composition, and that the results are robust and do not rely on unique properties of immunizations. Further explorations of these results support the proposition that the increase in immunization is rooted in its club good nature, as those with a greater share of the cost that is non-rival exhibit greater impacts of cohort size. The magnitudes of

the estimated effect sizes are reasonably large. If a child with one female age-mate has a 70% chance of immunization, she has a 3% chance of dying from a vaccine-preventable disease by age 5. If that child had three more age-mates, immunization increases to 79%; if they are all boys, 83%. In that case, the chance of vaccine-preventable death drops to 1.7% – a 44% reduction over the baseline case. This suggests that, at least with respect to certain investments, children with larger cohorts may be better off.

On a final note: the effects presented here are estimated based on exogenous non-sibling age-mates, and as such, do not warrant direct conclusions regarding the effects of fertility on child investment. Nonetheless, many programs assume that reducing fertility and increasing birth spacing (thereby reducing cohort sizes) would automatically improve child outcomes; an assumption that ignores potential club goods benefits. I leave open for future work the collection of adequate data that would allow a more direct test of how sibships affect investment in club goods.

# References


Aaby, P., Bukh, J., Lisse, I. M., and Smits, A. J. (1984). Overcrowding and intensive exposure as determinants of measles mortality. *American Journal of Epidemiology*, 120(1):49–63.

Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxf. Econ. Pap.*, 58(3):450–474.

Angrist, J., Lavy, V., and Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28(4):773–824.

Asfaw, A., Lamanna, F., and Klasen, S. (2010). Gender gap in parents' financing strategy for hospitalization of their children: Evidence from India. *Health Economics*, 19:265–279.

Barten, A. P. (1964). *Econometric Analysis for National Economic Planning*, chapter Family Composition, Prices and Expenditure Patterns. Butterworth.

Becker, G. and Lewis, G. (1973). On the interaction between the quantity and quality of children. *The Journal of Political Economy*, 81(2):S279–S288.

Becker, G. S. (1960). *Demographic and Economic Change in Developed Countries*, chapter An Economic Analysis of Fertility. Princeton University Press, Princeton, N.J.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2):669–700.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics*, 122(1):409–439.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2010). Small family, smart family? *Journal of Human Resources*, 45(1):33–58.

Bommier, A. and Lambert, S. (2004). Human capital investments and family composition. *Applied Economics Letters*, 11(3):193–196.

Caceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *The Journal of Human Resources*, 41(4):738–754.

Case, A. and Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, 116(3):499–532.

Clemens, M. (2009). Skill flow: A fundamental reconsideration of skilled-worker mobility and development. Center for Global Development Working Paper No. 180.

Dayioglu, M., Kirdar, M. G., and Tansel, A. (2009). Impact of sibship size, birth order and sex composition on school enrollment in urban Turkey. *Oxford Bulletin of Economics and Statistics*, 71(3):399–426.

Duraisamy, P. and Duraisamy, M. (1995). Determinants of investment in health of boys and girls: Evidence from rural households of Tamil Nadu, India. *Indian Economic Review*, 30(1):51–68.

Edmonds, E. V. (2006). Understanding sibling differences in child labor. *Journal of Population Economics*, 19:795–821.

Ejrnaes, M. and Portner, C. (2004). Birth order and the intrahousehold allocation of time and education. *The Review of Economics and Statistics*, 86(4):1008–1019.

Engel, E. (1895). Die lebenskosten belgischer arbeiter-familien fruher und jetzt. *Bulletin de l'Institut International de Statistique*, (9):1–129.

Filmer, D., Friedman, J., and Schady, N. (2009). Development, modernization, and childbearing: The role of family sex composition. *The World Bank Economic Review*, 23(3):371–398.

Garg, A. and Morduch, J. (1998). Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics*, 11:471–493.

Gong, X., van Soest, A., and Zhang, P. (2005). The effects of the gender of children on expenditure patterns in rural China: A semiparametric analysis. *Journal of Applied Econometrics*, 20(4):509–27.

Gustafsson, B. and Kjulin, U. (1994). Time use in child care and housework and the total cost of children. *Journal of Population Economics*, 7(3):287–306.

Joshi, S. and Schultz, T. P. (forthcoming). Family planning and women's and children's health: Consequences of an outreach program in Matlab, Bangladesh. *Demography*, (951).

Kambhampati, U. S. and Rajan, R. (2008). The 'nowhere' children: Patriarchy and the role of girls in India's rural economy. *Development Studies*, 44(9):1309–1341.

Kaufmann, C. and Briegel, H. (2004). Flight performance of the malaria vectors anaopheles gambiae and anopheles atroparvus. *Journal of Vector Ecology*, 29(1):140–53.

Lundberg, S. (2005). Sons, daughters, and parental behavior. *Oxford Review of Economic Policy*, 21(3):340–56.

Lynam, P. F., Smith, T., and Dwyer, J. (1994). Client flow analysis: A practical management technique for outpatient clinic settings. *International Journal for Quality in Health Care*, 6(2):179–186.

Makepeace, G. and Pal, S. (2008). Understanding the effects of siblings on child mortality: evidence from India. *Journal of Population Economics*, 21:877–902.

Mishra, V., Roy, T. K., and Retherford, R. D. (2004). Sex differentials in childhood feeding, health care, and nutritional status in India. *Population and Development Review*, 30(2):269–295.

Morduch, J. (2000). Sibling rivalry in Africa. *The American Economic Review*, 90(2):405–409.

Newman, R. D., Gloyd, S., Nyangezi, J. M., Machobo, F., and Muiser, J. (1998). Satisfaction with outpatient health care services in manica province, mozambique. *Health Policy and Planning*, 13(2):174–180.

Ono, H. (2004). Are sons and daughters substitutable? Allocation of family resources in contemporary Japan. *Journal of the Japanese and International Economies*, 18(2):143–60.

Ota, M. and Moffatt, P. G. (2007). The within-household schooling decision: A study of children in rural Andhra Pradesh. *Journal of Population Economics*, 20:223–239.

Pande, R. (2003). Selective gender differences in childhood nutrition and immunization in rural India: The role of siblings. *Demography*, 40(3):395–418.

Parish, W. and Willis, R. J. (1993). Daughters, education, and family budgets: Taiwan experiences. *The Journal of Human Resources*, 28(4):863–898.

Qian, N. (2009). Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural China. NBER Working Paper 14973.

Rammohan, A. and Dancer, D. (2008). Gender differences in intrahousehold schooling outcomes: The role of sibling characteristics and birth-order effects. *Education Economics*, 16(2):111–126.

Rose, E. (2000). Gender bias, credit constraints and time allocation in rural India. *The Economic Journal*, 110(465):738–758.

Rosenzweig, M. R. and Schultz, P. T. (1987). Fertility and investments in human capital : Estimates of the consequence of imperfect fertility control in malaysia. *Journal of Econometrics*, 36(1-2):163–184.

Rosenzweig, M. R. and Wolpin, K. I. (1980). Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy*, 88(2):328.

Rosenzweig, M. R. and Zhang, J. (2006). Do population control policies induce more human capital investment? Twins, birthweight, and China's one child policy. Yale University Economic Growth Center Discussion Paper No. 933.

Rosenzweig, M. R. and Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies*, 76(3):1149–1174.

SaveTheChildren (2012). Equitable health financing: Senegal. Online at www.equitablehealthfinancing.org/countries/senegal.

SEATS (2000). Mainstreaming quality improvement in family planning and reproductive health services delivery: Context and case studies. Technical Report PNACH 661, USAID. SEATS is the Family Planning Service Expansion and Technical Support Project.

Tanser, F., Gijsbertsen, B., and Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural south africa: An exploration using a geographical information system. *Social Science & Medicine*, 63(3):691–705.

Umar, N. A., Hajara, M. J., and Khalifa, M. (2010). Reduction of client waiting time using task shifting in an anti-retroviral xlinic at specialist hospital bauchi, nigeria. *Journal of Public Health in Africa*, 1(8).

Yamaguchi, K. (1989). A formal theory for male-preferring stopping rules of childbearing: Sex differences in birth order and in the number of siblings. *Demography*, 26(3):451–465.

Yount, K. M. (2003). Gender bias in the allocation of curative health care in Minia, Egypt. *Population Research and Policy Review*, 22(3):267–99.

# Appendix

## A. Are co-resident families' fertility decisions correlated?

In this section I present evidence that non-polygynous multi-family households are not coordinating on fertility timing. This is a key assumption underlying this analysis, as the birth timing of co-resident children is assumed to be reasonably exogenous, once time-invariant household characteristics have been controlled using household fixed-effects. If co-resident families were coordinating birth timing, that would suggest that conception by one woman would predict either the conception or non-conception of another woman in the household within 12 months or less (depending on whether the household preferred co-incident or spaced births).

I create a woman-by-month panel of fertility outcomes for rural, ever-married women based on the full birth histories collected in the Senegal DHS. For each woman in each month, the panel contains her age, parity, ever-married status, and whether she conceived in that month (calculated as 9 months prior to observed births). For each woman-month observation, I create an indicator for whether another woman in her household conceived within the 6 or 12 months prior to that month. Using a linear probability model, I regress whether a woman conceives in a given month on an indicator for recent conception by another woman in her household, controlling for age, parity, and ever-married status.

Fifty-three percent of women in multi-family households are not in a polygynous marriage. The results in table A.1 show that for these women, whether she conceives in a given month is not significantly predicted by whether another women in her household conceived in the 6 or 12 months prior. Despite very large sample sizes, these estimates fail to provide evidence that multi-family households are coordinating birth timing to achieve either co-incident or spaced births.

For women in polygynous marriages, I do find evidence of fertility coordination. These women are significantly more likely to conceive if a co-wife has conceived within the past 6 or 12 months (9% more likely and 5% more likely, respectively). Importantly, I also find that this predictive effect is no longer significant when household fixed effects are included.[23] That is, fertility correlations are driven by the household commonalities (e.g. the husband) that are

[23] I also included household fixed effects in the non-polygynous estimations, and the results are not different from those presented.

controlled by the household fixed effect.

## B. Parental autonomy within the household

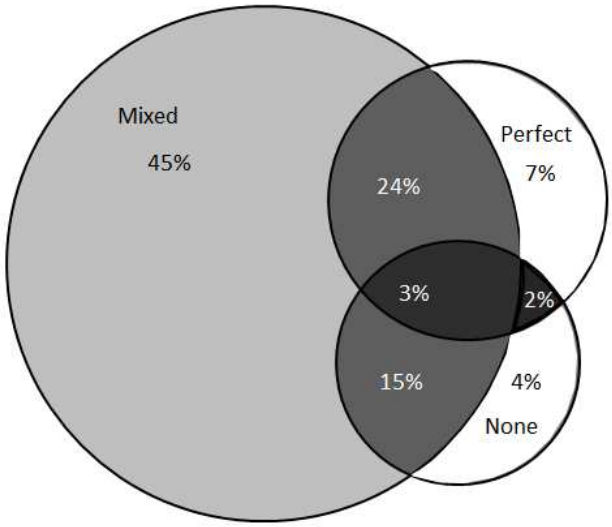
As an indicator of the ability of parents to make child investment decisions independently from the household, I examine differences in immunization records within a household. I focus on rural households that have multiple mothers of children under age 5. I categorize mothers as (i) having a perfect record (i.e. all children are current on all immunizations), (ii) none (no children have any immunizations), or (iii) mixed (having some but not all immunizations for which her children are due). Perfectly homogenous households contain women who all have perfect or all have "none" records. These households comprise 7% and 4% of all rural households with multiple mothers, respectively. Households with some intermediate or unknown level of homogeneity are those in which all mothers have mixed records (45%). Households with the greatest heterogeneity of immunizations across mothers are those where a "perfect record" mother or a "none" mother co-resides with "mixed" mothers or her opposite in terms on immunizations. These households combine to comprise 44% of all households. Figure B.1 shows these categorizations graphically, with the white areas being homogenous households, and successively darker shades indicating increasing levels of heterogeneity. The highest level of heterogeneity is when a "perfect record" mother co-resides with a "none" mother (5% of households).

Table A.1: Non-polygynous multi-family households do not coordinate fertility timing

| | Non-polygynous | | Polygynous | | | |
|---|---|---|---|---|---|---|
| | 6mo | 12mo | 6mo | 12mo | 6mo | 12mo |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Co-resident mother conceived in previous months | .00044 | -.00046 | .00146*** | .00079** | .00072 | -.00020 |
| | (.00050) | (.00037) | (.00048) | (.00034) | (.00048) | (.00034) |
| HH Fixed effects | No | No | No | No | Yes | Yes |
| Observations | 461788 | 461788 | 646716 | 646716 | 646716 | 646716 |
| $R^2$ | .007 | .007 | .005 | .005 | .005 | .005 |
| Mean Dependent Var | .01368 | .01368 | .01608 | .01608 | .01608 | .01608 |

Estimations at the woman-month level. Dependent variable is Y/N woman conceived in a given month. Estimations control for woman's age, parity and ever-married status specific to the month.

Figure B.1: Shares of Rural Multiple-Mother Households, by immunization record of resident mothers




Source: Author's calculations based on DHS data.
Note: "Perfect" denotes that the household contains a mother with a perfect immunization record. "None" denotes that the household contains a mother whose children have no immunizations. "Mixed" denotes that the household contains a mother who has gotten some but not all of the immunizations for which her child(ren) is (are) due. Darker shades indicate higher levels of heterogeneity within the household.

# C. First stage estimations from 2SLS

Table C.2: Predictive power of non-siblings for full age-cohort characteristics
Dependent variable shown in column header

| | Full Sample | Restricted Sample | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | $N_{ij}$ | $N_{ij}$ | $Z_{ij}$ |
| Numb. non-sib age-mates | .861*** | .851*** | -.004* |
| | (.011) | (.013) | (.002) |
| Male share of non-sib age-mates | | -.031 | .805*** |
| | | (.027) | (.010) |
| Child characteristics | 3 | 3 | 3 |
| Vaccine fixed effects | 9 | 9 | 9 |
| Household fixed effects | 4415 | 1639 | 1639 |
| Observations | 92643 | 50552 | 50552 |
| $R^2$ | .684 | .705 | .858 |
| Overall F-stat | 646 | 438 | 798 |

Notes: Cols 1 and 2 estimate eq. 4 and col 3 estimates eq. 5. Cols 2 and 3 employ only children with non-zero cohort size, in order to include the indicator of cohort gender composition. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects. Standard errors are show in parentheses, clustered at the household level.

# D. Main Results: Disaggregated by survey round

Table D.3: Main results by survey round
Dependent variable is Y/N immunization received

| | 1993 | | 2005 | |
|---|---|---|---|---|
| | Full | NZC Sample | Full | NZC Sample |
| | (1) | (2) | (3) | (4) |
| Size of age-cohort | .024** | .031*** | .031*** | .034*** |
| | (.011) | (.011) | (.007) | (.008) |
| Male share of age-cohort | | .087** | | .024 |
| | | (.038) | | (.019) |
| Observations | 30783 | 17545 | 61856 | 33007 |
| $R^2$ | .072 | .069 | .159 | .152 |

Notes: Cols 1 and 2 employ data from 1993 only. Cols 3 and 4 employ data from 2005 only. Cols 2 and 4 employ only children with non-zero cohort size, in order to include the indicator of cohort gender composition. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects. Standard errors are show in parentheses, clustered at the household level.

## Tables

Table 1: Senegalese National Immunization Schedule & Coverage

| | | Of children due, % that have received | | Sample Size | | |
|---|---|---|---|---|---|---|
| Vaccine Dose | Due at age | 1993 | 2005 | 1993 | 2005 | Total |
| Tuberculosis | Birth | 74% | 87% | 3,687 | 6,649 | 10,336 |
| Oral Polio Vaccine 0 | Birth | .. | 39% | 0 | 6,642 | 6,642 |
| Diptheria, Pertussis, Tetanus 1 | 6 weeks | 66% | 87% | 3,577 | 6,486 | 10,063 |
| Oral Polio Vaccine 1 | 6 weeks | 67% | 88% | 3,577 | 6,489 | 10,066 |
| Diptheria, Pertussis, Tetanus 2 | 10 weeks | 59% | 82% | 3,494 | 6,327 | 9,821 |
| Oral Polio Vaccine 2 | 10 weeks | 59% | 79% | 3,494 | 6,331 | 9,825 |
| Diptheria, Pertussis, Tetanus 3 | 14 weeks | 49% | 71% | 3,412 | 6,172 | 9,584 |
| Oral Polio Vaccine 3 | 14 weeks | 49% | 65% | 3,412 | 6,176 | 9,588 |
| Measles | 9 months | 58% | 76% | 3,071 | 5,304 | 8,375 |
| Yellow Fever | 9 months | 57% | 76% | 3,063 | 5,280 | 8,343 |
| Total | | 60% | 75% | 30,787 | 61,856 | 92,643 |

Notes: Rural sample only. Vaccines for which a child is not yet due are excluded. Information regarding Oral Polio Vaccine-0 was not collected in 1993.

Table 2: Sample Size

| Rural Sample | 1993 | 2005 | Total |
|---|---|---|---|
| Households | 1,444 | 3,046 | 4,490 |
| Mothers | 2,325 | 4,674 | 6,999 |
| Children U5 | 3,746 | 7,364 | 11,110 |
| Children U5 x vaccines | 33,696 | 73,640 | 107,336 |
| Children U5 x vaccines due | 31,357 | 68,344 | 103,408 |
| Children U5 x vaccines reported | 30,787 | 61,856 | 92,643 |
| CU5 with NSAM x vaccines reported | 17,545 | 33,057 | 50,602 |

Notes: “CU5” is children under age 5; “NSAM” is non-sibling age-mates.

Table 3: Summary Statistics for Rural Sample

| | Proportion or Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| **Households** $N = 5,907$ | | | | |
| Number of members | 9.6 | 6.08 | 1 | 80 |
| Number of offspring of HH head | 3.7 | 2.88 | 0 | 22 |
| Number of adopted/foster members | 1.1 | 2.79 | 0 | 38 |
| Has children U5 | 0.76 | | | |
| Has more than 2 adults | 0.82 | | | |
| of which, add'l adults are unrelated | 0.34 | | | |
| **Households with children U5** $N = 4,490$ | | | | |
| Children U5 | 2.5 | 1.65 | 1 | 18 |
| Mothers of children U5 | 1.5 | 0.91 | 0 | 10 |
| Has children U5 with age-mates | 0.53 | | 0 | 1 |
| Has variation in number of age-mates across CU5 in household | 0.66 | | 0 | 1 |
| **Mothers of children U5** $N = 6,999$ | | | | |
| Age | 29 | 7.52 | 15 | 49 |
| Completed primary school | 0.13 | | 0 | 1 |
| Number of children | 4.3 | 2.72 | 1 | 15 |
| Number of children U5 | 1.6 | 0.64 | 1 | 5 |
| Has coresident mother(s) of CU5 | 0.56 | | 0 | 1 |
| Has coresident mother(s) that are not co-wives | 0.31 | | 0 | 1 |
| Has coresident mother that is co-wife | 0.14 | | 0 | 1 |
| **Children U5** $N = 11,110$ | | | | |
| Percent of due vaccines received | 0.63 | 0.38 | 0 | 1 |
| Has any age-mates in household | 0.68 | 0.47 | 0 | 1 |
| if so, number of age-mates | 2.31 | 1.65 | 1 | 15 |
| if so, non-sibling age-mates | 1.82 | 1.69 | 0 | 14 |
| Share of age-mates that are boys | 0.51 | 0.4 | 0 | 1 |

Notes: "CU5" is children under age 5. Age-mates are other children in the household within 24 months of one's age. For 20% of women with co-resident mothers (11% of all rural women), determination of co-wife status is not possible, see notes to Table 4.

Table 4: Relations of Co-resident Mothers

| Woman's relation to household head | Women in households with Only one mother | CRMs that are not co-wives | CRM with co-wife status unknown | Total |
|---|---|---|---|---|
| Self | 6% | 1% | 7% | 4% |
| Wife | 55% | 24% | .. | 42% |
| Daughter | 9% | 10% | 10% | 8% |
| **Daughter-in-law** | **13%** | **23%** | **28%** | 17% |
| **Foster/Adopted daughter** | **5%** | **20%** | **19%** | 11% |
| Other relative | 9% | 15% | 31% | 15% |
| Not related | 2% | 5% | 5% | 3% |

Notes: The table shows the distribution of relations to the household head for each household type, according to presence and/or co-wife status of co-resident mothers (CRMs). The distribution is not shown for CRMs that are co-wives, as the ability to identify them is contingent on relation to household head, as described below. CRMs that are co-wives are however included in the total. The following method was employed to determine whether CRMs are co-wives:

- For those responding that they are not in a polygynous union, clearly CRMs are not co-wives.
- For those that are in a polygynous union:
    1. If the respondent is listed as wife of the household head, co-wife status of CRMs is determined by whether any CRM is also listed as wife of head;
    2. For female heads of household, CRMs listed as "co-spouse" are that, and vice versa;
    3. For those not listed as wife, head, or co-spouse, whether CRMs are co-wives cannot be determined (11% of sample). However, it is of note that for women in polygynous unions listed as wife of head, 50% of the time her CRMs are not co-wives (i.e. co-wives live elsewhere); therefore for the 11% of women for whom co-wife status of CRMs is indeterminate, it is likely that half will have CRMs that are not co-wives.

Table 5: Effect of Age-cohort on Immunization
Dependent variable: Y/N vaccine received

| | Endogenous Siblings | | Exogenous Non-siblings | | Full Age-cohort | |
|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) IV | (6) IV |
| Size of age-cohort | -.017† | -.031 | .025*** | .028*** | .029*** | .033*** |
| | (.009) | (.021) | (.005) | (.006) | (.006) | (.007) |
| Male share of age-cohort | | .004 | | .036** | | .046** |
| | | (.022) | | (.015) | | (.018) |
| Child characteristics | 3 | 3 | 3 | 3 | 3 | 3 |
| Vaccine fixed effects | 9 | 9 | 9 | 9 | 9 | 9 |
| Household fixed effects | 4415 | 1506 | 4415 | 1639 | 4415 | 1639 |
| Observations | 92643 | 26539 | 92643 | 50552 | 92643 | 50552 |
| $R^2$ | .128 | .134 | .129 | .123 | .128 | .120 |

Notes: Column headers indicate those included in "age-cohort." Cols 1 and 2 estimate eq. 3; cols 3 and 4 estimate eq. 2; cols 5 and 6 are eq. 6, the second stage of 2SLS. Columns 1, 3, and 5 employ the full rural sample. Columns 2, 4 and 6 employ the sub-sample that has a non-zero age-cohort. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects. Standard errors are in parentheses, clustered at the household level. *** Indicates significance at the .001 level; ** at the .01 level, * at the .05 level, and † at the .10 level.

Table 6: Heterogeneous Effects; 2SLS
Dependent variable: Y/N vaccine received

| | With clinic in village | | Full sample | | NZC sample | |
|---|---|---|---|---|---|---|
| | Full | NZC Sample | NoEduc | AnyEduc | NoEduc | AnyEduc |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Size of age-cohort | .014 | .018 | .026*** | .061*** | .026*** | .069*** |
| | (.014) | (.015) | (.006) | (.016) | (.007) | (.018) |
| Male share of age-cohort | | .094 | | | .032 | .072 |
| | | (.058) | | | (.019) | (.047) |
| Child characteristics | 3 | 3 | 3 | 3 | 3 | 3 |
| Vaccine fixed effects | 9 | 9 | 9 | 9 | 9 | 9 |
| Household fixed effects | 409 | 173 | 3919 | 1573 | 764 | 366 |
| Observations | 8505 | 5112 | 77549 | 11643 | 42603 | 5864 |
| $R^2$ | .089 | .084 | .144 | .151 | .134 | .155 |

Notes: Estimations shown are variations on eq. 6; the second stage of 2SLS. Columns 1 & 2 employ rural households that have a health clinic in their village (information available in 1993 data only). Cols 3 - 6 show sub-samples according to mother's education. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects. Standard errors are in parentheses, clustered at the household level. *** Indicates significance at the .001 level; ** at the .01 level, * at the .05 level, and † at the .10 level.

Table 7: Alternative Specifications; 2SLS
Dependent variable: Y/N vaccine received

| | (1) | (2) | (3) |
|---|---|---|---|
| Size of age-cohort | .049*** | .036*** | .025** |
| | (.013) | (.007) | (.009) |
| (Size)$^2$ | -.002† | | |
| | (.001) | | |
| Size x Male | | -.005 | |
| | | (.005) | |
| Size x Male share | | | .017 |
| | | | (.015) |
| Male share of age-cohort | .091 | .058* | .020 |
| | (.064) | (.026) | (.028) |
| (Male share)$^2$ | -.045 | | |
| | (.062) | | |
| Male share x Male | | -.017 | |
| | | (.037) | |
| Child characteristics | 3 | 3 | 3 |
| Vaccine FE | 9 | 9 | 9 |
| Household FE | 1639 | 1639 | 1639 |
| Observations | 50602 | 50602 | 50602 |
| $R^2$ | .12 | .12 | .12 |
| $\chi^2$(Male share, (Male share)$^2$) | 7.16 | | |
| $\chi^2$(Male share, Size x Male share) | | | 7.21 |

Notes: Estimations shown are variations on eq. 6; the second stage of 2SLS. All columns show rural sample of children with any exogenous age-mates (NZC sample). "Male share" is the proportion of one's cohort that is male. "Male" indicates that child *i* is male. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects. Standard errors are in parentheses, clustered at the household level. *** Indicates significance at the .001 level; ** at the .01 level, * at the .05 level, and † at the .10 level.

Table 8: Robustness Checks
Dependent variable: Y/N vaccine received, except as noted (col. 1)

| | Child-level DV: Deceased | Excluding potential violators | Excluding vaccs. for human transmission | Age-mate is within… | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | 36mo. | 30mo. | 18mo. | 12mo. | 12mo. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Size of age-cohort | -.014* | .027*** | .028* | .032*** | .038*** | .023** | .014 | .014* |
| | (.006) | (.007) | (.011) | (.007) | (.007) | (.007) | (.010) | (.007) |
| Male share of age-cohort | -.010 | .044* | .051 | .057† | .062* | .035* | .039** | |
| | (.014) | (.019) | (.031) | (.034) | (.025) | (.016) | (.015) | |
| Child Characteristics | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| Vaccine FE | 0 | 9 | 0 | 9 | 9 | 9 | 9 | 9 |
| Household FE | 1649 | 1520 | 1338 | 1773 | 1712 | 1549 | 1367 | 4412 |
| Obs. | 6048 | 44953 | 4327 | 54781 | 53001 | 47010 | 39937 | 92639 |
| $R^2$ | .011 | .121 | .024 | .12 | .118 | .122 | .124 | .128 |

Notes: Col. 1 is estimated at the child level and the dependent variable is Y/N deceased. All other estimations shown are variations on eq. 6, the second stage of 2SLS, and the dependent variable is Y/N vaccine received. All columns except the last use the sample of rural children with a non-zero exogenous cohort (NZC), with exclusions as noted. Estimates based on the full sample do not differ from these, with the exception of the 12 month definition. The final column shows this estimate for the full sample, where the effect is significantly different from zero. The lack of precision in the 12mo-NZC sample is likely due to the reduced sample size. All estimations include controls for child gender, birth order, and household birth order, as well as household and vaccine fixed effects (except as noted). Standard errors are in parentheses, clustered at the household level. *** Indicates significance at the .001 level; ** at the .01 level, * at the .05 level, and † at the .10 level.