

Family Size and Educational Investments in Children: Evidence from Private Tutoring Expenditures in South Korea

Changhui Kang
Department of Economics
National University of Singapore
1 Arts Link, Singapore 117570
Singapore
E-mail: ecskch@nus.edu.sg
Phone: +65-6516-6830
Fax: +65-6775-2646

Abstract

Relying on private tutoring expenditures in South Korea, this paper examines whether large family size has a strong negative impact on educational investments in children. In the face of difficulties in finding an exogenous IV for family size, the paper employs a nonparametric bounding method. Our primary finding is that quantity-quality trade-offs in educational investments function in a way that varies by the sex of the child. While there is a non-negligible negative effect of large family size on educational investments for girls, there is little if any impact on those for boys. Son preferences traditionally shown by Korean parents seem to underlie such empirical findings.

JEL Classification: I20, C30

Keywords : Family size, Private Tutoring, South Korea, Nonparametric bounds

Family Size and Educational Investments in Children: Evidence from Private Tutoring Expenditures in South Korea

1 Introduction

One of the well-established empirical regularities in economics of family is that children in large families receive small educational investments and show poor educational attainment (Blake, 1989; Hanushek, 1992; Hauser and Sewell, 1986; Rosenzweig and Wolpin, 1980). In light of a potential link between population growth and economic growth, such a tradeoff between quantity and quality of children is often promoted as strong evidence that small family and slow population growth lead to economic development (de la Croix and Doepke, 2003; Galor and Weil, 2000; Moav, 2005). Recently, however, researchers raise questions about whether the empirical negative relationship between family size and children's education implies a true causal effect of family size on children (Angrist et al., 2006; Caceres-Delpiano, 2006; Conley and Glauber, 2006; Guo and Van Wey, 1999). For example, parents with higher socio-economic status (SES) or cognitive abilities may have smaller families and invest more in education of each child than parents with lower SES or cognitive abilities; even in the absence of true causality, this may yield a strong negative correlation between family size and educational investments (and attainment) of a child.

In order to purge spurious correlations, several recent empirical studies rely on instrumental variables (IV) to identify exogenous changes in family size. Three types of IVs are in frequent use in empirical analysis. One group of studies use multiple birth (e.g., twin-birth) as a source of exogenous variation in family size (Angrist et al., 2006; Black et al., 2005; Black et al., 2007b; Caceres-Delpiano, 2006; Rosenzweig and Wolpin, 1980; Rosenzweig and Zhang, 2006). Another group of studies employ sibling sex composition as an IV for family size (Angrist et al., 2006; Black et al., 2007b; Conley and Glauber, 2006; Lee, 2007). The other strand relies on institutional changes that give rise to changes in fertility decisions of parents (e.g., China's one-child policy) (Li and Zhang, 2007; Liu, 2007; Qian, 2006).

In spite of alleged exogeneity of each of such IVs, there are several doubts as to whether each of these IVs are truly exogenous and whether the studies deliver the relevant empirical evidence

that they desire to deliver. First, it is well known that twin-birth may be confounded with zero birth-spacing between siblings. If zero birth-spacing has independent effects on raising non-twin children either via physical and financial stress for parents or via scale economies (e.g., Downey et al., 1999), the ultimate empirical evidence about family size effects based on twins is not perfectly clear. In addition, rates of low birth weight and infant mortality for twins are much higher than those of singleton births (Almond et al., 2005; Black et al., 2007a; Rosenzweig and Zhang, 2006). If parents of twins allocate family resources across siblings in a different way than those of singletons do, the impacts of large family size on education of non-twin siblings may differ between twin and non-twin families (Rosenzweig and Zhang, 2006).

Second, there are also considerable doubts as to whether sibling sex composition affects children's educational investments and attainment via changes in family size alone. Rosenzweig and Wolpin (2000) suggest there are cost savings associated with raising same sex siblings; such savings may independently lead to greater educational investments and outcomes of children. Liu (2007) argues that if sex mix of siblings is greatly appreciated in parents' preferences, parents may invest *more* in education of each child when they have large size of mixed-sex siblings: for example, parents may *increase* educational investments in two first sons with an arrival of one younger daughter, if mixed sex of siblings reinforces parents' satisfaction for increased quality of each child at the expense of reduced consumption. In addition, Ananat and Michaels (2007) and Dahl and Moretti (2004) show effects of sibling sex composition on marriage outcomes, which may have independent effects on educational investments for children.

Third, institutional changes can also be called into question if parents respond to exogenous policy changes by moving across different regions and/or if fertility-related policies are introduced in consideration of temporal and regional patterns of fertility (Moffitt, 2005; Rosenzweig and Zhang, 2006). In sum, there is a considerable amount of difficulties to find truly exogenous (ex ante as well as ex post) IVs in an analysis of causal effects of family size.

In the face of such difficulties, the current study examines the effect of family size on educational investments for children by employing an alternative strategy that is recently gaining popularity in empirical analysis—a nonparametric bounding method. Nonparametric bounds analysis was first introduced in economics by Manski (1990) and further developed in Manski (1997), Manski and Pepper (2000) and Hotz et al. (1997). Some recent examples of this method include Blundell et al. (2007), Gerfin and Schellhorn (2006), Gonzalez (2005), Kreider

and Pepper (2007), Lechner (1999), Manski and Nagin (1998), and Pepper (2000) among others.

The basic idea of the nonparametric bounding method is that instead of obtaining point estimates that often rely on questionable assumptions, one may calculate lower and upper bounds of the treatment effect given a few weaker assumptions. A unique advantage of this approach in an estimation of effects of family size is that a potential IV does not have to be fully exogenous. To the extent that the IV is monotonically (either positively or negatively) related with the outcome variable, the method, in combination of some other assumptions, is able to draw fairly tight bounds of the causal effect (Manski and Pepper, 2000). Provided that the bounds are sufficiently narrow and informative to locate the causal effect, we interpret that the magnitude of the true effect is somewhere between the estimated lower and upper bounds.

We apply such a nonparametric bounding method to the patterns of private tutoring expenditures for school-age children in South Korea. South Korea offers an interesting example in the current context: the country has widespread and large-scale markets for private tutoring, and monetary expenditures on tutoring can serve as a good proxy for educational investments for a child, which is rarely available elsewhere by typical household surveys. Moreover, because there are virtually no private secondary schools that are independent of the government's control in Korea¹, tutoring expenditures may not be confounded with costs of attending a private school. In Korea, private tutoring is largely performed as a supplementary learning on top of the formal (public) school education. As a result, variation in expenditure on private tutoring for children is highly likely to measure that in total monetary educational investment made by the parents.

To the extent that economic theories on quantity-quality trade-offs are based on parental choices of educational resource allocation across siblings (Becker and Lewis, 1973; Becker and Tomes, 1976), monetary expenditures on private tutoring are arguably more appropriate to test the theoretical predictions than, for example, private school attendance and educational attainments used in other studies (Caceres-Delpiano, 2006; Conley and Glauber, 2006). Lee (2007) has first used tutoring expenditures of Korean parents as a proxy for educational investments in an analysis of family size effects. The current paper shares with Lee's study the measure of educational investments and first-born sex as an (monotone) IV for family size. But it differs

¹In Korea private middle and high schools are little different from public schools with respect to school administration, curriculum and student placement, because they are heavily subsidized and controlled by the government. For an overview of secondary education and private tutoring in South Korea, see Kang (2007), Kim and Lee (2001) and OECD (1998).

from Lee’s study in at least three respects. First, the current study employs a different empirical approach—nonparametric bounding method—in order to consider potential endogeneity of first-born sex as an IV. Second, as Lee’s data contain information only on overall expenditures on private tutoring for *all* children in the family, his study fails to control for a child’s birth order in a study of family size effects. Black et al. (2005), for example, underscore the importance of controlling for birth order in examining family size effects. In contrast, the current study relies on a data set that has information on private tutoring expenditures for each child in the family, which enables us to control for birth order. Third, the current paper examines private tutoring of more recent cohorts (after 2000) of children than used by Lee’s paper, which looks at private tutoring in the mid-1990s before private tutoring markets grew rapidly in Korea at the turn of the century.

Employing the nonparametric bounding method and private tutoring expenditures in South Korea, the current paper shows that large (small) family size has a strong negative (positive) impact on educational investments for girls but little impact on those for boys. Namely, quantity-quality trade-offs in educational investments function in a way that varies by the sex of the child. Son preferences traditionally shown by Korean parents seem to underlie such empirical findings. We discover that second-born girls suffer more in education from increased family size if a younger brother is born to the family than if a younger sister is born; in contrast, second-born boys suffer as little—or slightly less due to reduced competition with other son(s)—when they have a younger sister as when they have a younger brother.

The rest of the paper is organized as follows. Section 2 describes the nonparametric bounding method. We explain the data in section 3 and the empirical results in section 4. Section 5 concludes the paper.

2 Empirical Framework

Following Angrist et al. (2006) and many others and for clear interpretation of the results, we restrict the analysis to second-born children (boys and girls, separately) of the family that has at least two children. Let us first define y_i as a natural log of an average monthly expenditure on private tutoring (in KRW 1,000) for second-born child i who attends school in grades 1 to 12. Let a treatment indicator T_i be equal to zero if the total number of i ’s siblings is equal to

one, and one if it is greater than one. Note that we set large (small) family size as a treatment (control). Finally, let an indicator D_i be equal to one if the first-born child of i 's family is a daughter, and 0 if it is a son. Below we employ D_i as a monotone IV for family size.²

Each child receives treatment $t \in T$. The response function $y_i(\cdot) : T \rightarrow Y$ maps treatments into outcomes. The realized outcome $y_i \equiv y_i(z)$ is the level of y for child i who actually receives treatment z . The latent outcome $y_i(t)$ ($t \neq z$) describes what level of educational investments would have been made for child i had he or she received treatment t . Of primary interest is the causal effect of having a large number of siblings in the family (i.e., more than one as opposed to only one) on monetary educational investments for a child. That is, $E[y_i(1) - y_i(0)]$.

Following Gonzalez (2005), Manski (1990) and Manski and Pepper (2000), in order to set up bounds for the treatment effect, we first decompose $E[y(t)]$ by

$$E[y(t)] = E[y|z = t]Pr(z = t) + E[y(t)|z \neq t]Pr(z \neq t) \quad (1)$$

To make bounds analysis feasible, let us suppose that y is bounded by $[K_0, K_1]$. Since the unobservable counterfactual $E[y(t)|z \neq t]$ is also bounded by $[K_0, K_1]$, we have the worst-case (WC) bounds of $E[y(t)]$ given by

$$\begin{aligned} & E[y|z = t]Pr(z = t) + K_0Pr(z \neq t) \\ & \leq E[y(t)] \leq \\ & E[y|z = t]Pr(z = t) + K_1Pr(z \neq t) \end{aligned} \quad (2)$$

In order to further tighten the bounds of $E[y(t)]$, a few assumptions can be invoked individually as well as jointly. The first assumption to be employed is monotone treatment response (MTR), which is specified as follows³:

$$t_1 < t_2 \longrightarrow y(t_1) \geq y(t_2) \quad (3)$$

²Son preference of parents and the presence of sex-selective abortions make it difficult to use sex composition of the first two children as an IV for family size in the Korean context. To the extent that sex selections are relatively rare for first-born children, the first-child's sex is a more appropriate IV for family size in Korea. See section 3 of Lee (2007) for further details.

³Note the reversion of the inequality on the right-hand side of the arrow compared with Gonzalez (2005) and Manski and Pepper (2000). Such a reversion is also shown in the monotone treatment selection (MTS) assumption below. Each of the nonparametric bounds under those two assumptions is adjusted in consideration of the reversed inequalities.

This assumption is drawn from economic theories that predict negative (at least, non-positive) impacts of large family size on educational investments in a child (Becker and Lewis, 1973; Becker and Tomes, 1976). Although some studies (e.g., Qian, 2006) document positive impacts of large family size on educational attainment of children, it is hard to imagine that, given a constrained budget, parents invest more on each child's education when the family size is large than when it is small. Under MTR, the following relationships hold:

$$\begin{aligned}
z < t &\longrightarrow K_0 \leq y(t) \leq y(z) \\
z = t &\longrightarrow y(t) = y(z) \\
z > t &\longrightarrow y(z) \leq y(t) \leq K_1
\end{aligned}$$

Hence, the MTR bounds of $E[y(t)]$ can be expressed by

$$\begin{aligned}
&E[y|z \geq t]Pr(z \geq t) + K_0Pr(z < t) \\
&\leq E[y(t)] \leq \\
&E[y|z \leq t]Pr(z \leq t) + K_1Pr(z > t)
\end{aligned} \tag{4}$$

The second assumption is monotone treatment selection (MTS), which is specified by:

$$t_1 < t_2 \longrightarrow E[y(t)|z = t_1] \geq E[y(t)|z = t_2] \tag{5}$$

This assumption supposes that sorting into treatment is not exogenous but monotone in the sense that the expected value of latent outcome $y(t)$ is greater (or equal) for those whose family size is small ($z = 0$) than for those whose family size is large ($z = 1$). For instance, parents of high socio-economic status are more likely to form a smaller family and invest more for each child's education than those of low socio-economic status. Such a scenario is likely to hold equally for second-born boys and girls. While it specifies a source of endogeneity in a conventional OLS method of examining the impacts of family size, the MTS assumption can make an important contribution to tightening the bounds of the true effect in a bounds analysis.

Under MTS, the following relationships hold:

$$\begin{aligned}
u < t &\longrightarrow E[y|z = t] \leq E[y(t)|z = u] \leq K_1 \\
u = t &\longrightarrow E[y(t)|z = u] = E[y|z = t] \\
u > t &\longrightarrow K_0 \leq E[y(t)|z = u] \leq E[y|z = t]
\end{aligned}$$

Hence, the MTS assumption yields the bounds of $E[y(t)]$ given by

$$\begin{aligned}
&E[y|z = t]Pr(z \leq t) + K_0Pr(z > t) \\
&\leq E[y(t)] \leq \\
&E[y|z = t]Pr(z \geq t) + K_1Pr(z < t)
\end{aligned} \tag{6}$$

The third assumption is the presence of a monotone instrumental variable (MIV) that is proposed by Manski and Pepper (2000). While an instrumental variable (IV) v satisfies mean-independence (i.e., $E[y(t)|v = u_1] = E[y(t)|v = u_2]$, $u_1 \neq u_2$), an MIV is assumed to satisfy the following mean-monotonicity:

$$u_1 < u_2 \longrightarrow E[y(t)|v = u_1] \leq E[y(t)|v = u_2] \tag{7}$$

$$\text{or } u_1 > u_2 \longrightarrow E[y(t)|v = u_1] \leq E[y(t)|v = u_2] \tag{8}$$

The MIV assumption in general supposes that, for a given family size, a second-born child in a daughter-first family ($D_i = 1$) receives greater (or smaller) educational investments than the counterpart in a son-first family ($D_i = 0$). If we examine the MIV assumption separately for boys and girls, different directions of mean-monotonicity may hold for each. Namely, equation (7) is more likely to be valid for second-born girls than equation (8) is, while equations (7) and (8) are likely to be equally valid for second-born boys.

First, if parents tend to prefer sons to daughters in a society such as South Korea (Park and Cho, 1995; Das Gupta et al., 2003), a second-born daughter can be as much or more advantaged in several dimensions including education in a family whose first-born is a daughter ($D_i = 1$) than in a family whose first-born is a son ($D_i = 0$). A second-born daughter is likely to receive as much or more educational investments in a daughter-daughter family than in a son-daughter family, because in the latter family a disproportionately greater share of educational resources

can be spent for a son.

Besides preferences of sibling sex composition, studies show that there are many forms of cost savings in raising same sex siblings. For example, Rosenzweig and Wolpin (2000) and Rosenzweig and Zhang (2006) find that parents of same sex siblings pay less money on clothing and others. Goux and Maurin (2005) report that same sex siblings tend to share the same room and live more often in overcrowded housing. If there are savings associated with siblings of same sex, they can be routed into educational investments in children. Second-born daughters in daughter-daughter families are more likely to benefit from such savings than those in son-daughter families.

The preceding MIV assumption that a second-born daughter is as much or more favored in a daughter-first family than in a son-first family requires that the statement is true not only in a two-children family ($t = 0$) but in a family with more than two children ($t = 1$). Such a requirement opens a possibility that the validity of MIV in fact depends on the sex composition of siblings in parity 3 or above. For example, even if a second-born girl of a two-children family is more favored in a daughter-first than in a son-first family, she may not be favored in the daughter-first family if she has third-born or fourth-born siblings. Nonetheless, to the extent that both son preference and cost savings affect educational investments of boys and girls in a similar degree, whether a parity-3 child is a boy or a girl does not affect the validity of equation (7) for second-born girls. Provided that each of son preference and cost savings concerns only the total number of sons and daughters in a family regardless of birth order, a second-born girl is as much or more advantaged in all four cases that follow: (1) she will be more favored in a daughter-daughter-daughter family than in a son-daughter-daughter family, because she benefits from the presence of less sons and more daughters in the former family; (2) she will be more favored in a daughter-daughter-son family than in a son-daughter-son family, and (3) in a daughter-daughter-daughter family than in a son-daughter-son family for the same reason; (4) she will be as much favored in a daughter-daughter-son family as in a son-daughter-daughter family, because sex composition of both families is the same.

However, MIV of equation (7) may be affected by the sex of a parity-4 sibling. For instance, a second-born girl can be less favored in a daughter-daughter-son-son family than in a son-daughter-son-son family, if a positive effect of cost savings under three sons exceeds a negative effect of son preference. Thus in the subsequent analysis we exclude the families that have more

than 3 children. Such a restriction, however, will not cause serious problems, since only a total of 49 observations (2.5 percent) are removed by the restriction. (If we include those children and maintain MIV of equation (7), the results are largely similar. They are available upon request.)

Second, while equation (7) seems more valid for second-born girls than equation (8), it is unclear which equation is more valid for second-born boys. While cost savings in raising same-sex siblings predict that second-born boys may receive greater educational investments in a son-son family than in a daughter-son family (i.e., equation (8)), son preference suggests that second-born sons may receive greater educational investments (due to less competition with other son for resources) in a daughter-son family than in a son-son family (i.e., equation (7)). Or, if son-preferring parents allocate a greater share of family budget to children's education in a son-son family than in a daughter-son family, son preference suggests greater educational investments in a second-born boy in the former family (i.e., equation (8)). (Equally plausible is that the level of educational investments for second-born boys is independent of the sex of the first-born, which supports the use of it as an IV rather than an MIV.)

As in the case of second-born girls, the validity of MIV for second-born boys may be affected by the sex composition of siblings in parity 3 or above. Nonetheless, as shown in Appendix, to the extent that either effect of son preference and cost savings is dominant over the other in the distribution of family educational resources among siblings, one MIV assumption (i.e., either equation (7) or (8)) holds both in two-children and three-children families. Namely, if son preference is dominant, equation (7) holds for a second-born boy both in two-children and three-children families; if cost savings from same-sex siblings are dominant, equation (8) holds both in two-children and three-children families.

Since MIV may also be compromised in families with more than three children, we exclude those families from the subsequent analysis. Such a restriction, however, will not cause serious problems, because only a total of 5 observations (0.2 percent) are removed by it. (If we include those observations and maintain MIV, the results are largely similar.) To the extent that there are equally valid two scenarios for second-born boys, we report two sets of MIV results. One set is based on equation (7); the other on (8). As shown below, however, both produce fairly similar empirical results.⁴

⁴Although Lee (2007, Table 4) shows macro evidence that the sex of first-born children is not artificially selected (say, via sex-selective abortions) by Korean parents, such evidence does not necessarily imply that the sex of the first-born is exogenously determined. Nonetheless, sex selection of the first-born by parents with a

The MIV assumption under (7) yields the bounds of $E[y(t)]$ given by⁵

$$\begin{aligned} & \sum_{u \in D} Pr(D = u) \{ \sup_{u_1 \leq u} [E(y|D = u_1, z = t)Pr(z = t|D = u_1) + K_0Pr(z \neq t|D = u_1)] \} \\ & \leq E[y(t)] \leq \\ & \sum_{u \in D} Pr(D = u) \{ \inf_{u_2 \geq u} [E(y|D = u_2, z = t)Pr(z = t|D = u_2) + K_1Pr(z \neq t|D = u_2)] \} \end{aligned} \quad (9)$$

Given each of the MTR, MTS and MIV assumptions, we impose multiple assumptions jointly in order to further tighten the bounds of $E[y(t)]$ and the treatment effect $E[y(1)] - E[y(0)]$. We below experiment with all possible combinations of the individual assumptions: MIV+MTR, MIV+MTS, MTR+MTS, and MIV+MTR+MTS.

Under MIV+MTR, the bounds of $E[y(t)]$ are given by

$$\begin{aligned} & \sum_{u \in D} Pr(D = u) \{ \sup_{u_1 \leq u} [E(y|D = u_1, z \geq t)Pr(z \geq t|D = u_1) + K_0Pr(z < t|D = u_1)] \} \\ & \leq E[y(t)] \leq \\ & \sum_{u \in D} Pr(D = u) \{ \inf_{u_2 \geq u} [E(y|D = u_2, z \leq t)Pr(z \leq t|D = u_2) + K_1Pr(z > t|D = u_2)] \} \end{aligned} \quad (10)$$

Under MIV+MTS, the bounds of $E[y(t)]$ are given by

$$\begin{aligned} & \sum_{u \in D} Pr(D = u) \{ \sup_{u_1 \leq u} [E(y|D = u_1, z = t)Pr(z \leq t|D = u_1) + K_0Pr(z > t|D = u_1)] \} \\ & \leq E[y(t)] \leq \\ & \sum_{u \in D} Pr(D = u) \{ \inf_{u_2 \geq u} [E(y|D = u_2, z = t)Pr(z \geq t|D = u_2) + K_1Pr(z < t|D = u_2)] \} \end{aligned} \quad (11)$$

Under MTR+MTS, the bounds of $E[y(t)]$ are given by

$$\begin{aligned} & \sum_{h > t} E(y|z = h)Pr(z = h) + E(y|z = t)Pr(z \leq t) \\ & \leq E[y(t)] \leq \\ & \sum_{h < t} E(y|z = h)Pr(z = h) + E(y|z = t)Pr(z \geq t) \end{aligned} \quad (12)$$

strong son preference does not invalidate the assumption that equation (7) is valid for second-born girls, but it strengthens the assumption because the presence of a son(s) has larger negative impacts on second-born girls under such a scenario. In addition, potential sex selection of the first-born child also strengthens equation (7) for second-born boys for the same reason, rendering it more valid than equation (8).

⁵In this section we will describe the MIV bounds under (7) whenever the MIV assumption is employed. The MIV bounds under (8), however, can be easily constructed if D is replaced by a new variable S where $S \equiv 1 - D$.

Finally, under MIV+MTR+MTS, the bounds of $E[y(t)]$ are given by

$$\begin{aligned}
& \sum_{u \in D} Pr(D = u) \cdot \\
& \quad \left\{ \sup_{u_1 \leq u} [\sum_{h > t} E(y|D = u_1, z = h)Pr(z = h|D = u_1) + E(y|D = u_1, z = t)Pr(z \leq t|D = u_1)] \right\} \\
& \leq E[y(t)] \leq \\
& \sum_{u \in D} Pr(D = u) \cdot \\
& \quad \left\{ \inf_{u_2 \geq u} [\sum_{h < t} E(y|D = u_2, z = h)Pr(z = h|D = u_2) + E(y|D = u_2, z = t)Pr(z \geq t|D = u_2)] \right\}
\end{aligned} \tag{13}$$

Given the bounds of $E[y(t)]$ under varying assumptions, the lower bound (LB) of average treatment effects (ATE), $E[y(1)] - E[y(0)]$, is calculated by the difference between the lower bound of $E[y(1)]$ and the upper bound of $E[y(0)]$; the upper bound (UB) of ATE is obtained by the difference between the upper bound of $E[y(1)]$ and the lower bound of $E[y(0)]$. Along with the bounds of $E[y(t)]$ and ATE are calculated bootstrap 5th and 95th percentiles of the lower and upper bounds, respectively. The interval between these percentiles shows a conservative 90% confidence interval for the estimated bounds. In addition, in order to take multiple observations for a same child into account, we generate each bootstrap sample by sampling the individual child (with replacement) first and using all expenditure observations of the child. The number of the bootstrap samples is 500.

There are some unusual cases where the estimated LB of $E[y(t)]$ exceeds its estimated UB due to the estimation bias for $E[\cdot]$. It arises especially when more than one assumption is jointly imposed and the bounds for $E[y(t)]$ become quite tight. In such cases the bounds of ATE can not be constructed by differencing between the estimated LB and UB of $E[y(1)]$ and $E[y(0)]$. If the estimated LB and UB of $E[y(t)]$ are reversed, we employ an alternative method of obtaining ATE, following an idea of Blundell et al. (2007, p.341). Under the null that the difference between the upper and lower bounds is zero, both the upper and lower bound estimates are consistent estimates of $E[y(t)]$. Thus one may choose the estimate of either the upper or lower bound as a consistent estimates for $E[y(t)]$; instead, we use a weighted combination of the upper and lower bounds, that is,

$$\widehat{E}[y(t)] = \alpha \widehat{E}^L[y(t)] + (1 - \alpha) \widehat{E}^U[y(t)]$$

where $\alpha \in [0, 1]$ is a weight, and $\widehat{E}^L[y(t)]$ ($\widehat{E}^U[y(t)]$) is the estimated lower (upper) bound of $E[y(t)]$. We calculate the bounds of ATE by setting $\alpha = 0.5$, while other values of α yield

qualitatively similar results. In tables of the results, the bounds of ATE that are obtained by such an alternative method are presented under a different row, while those calculated with the regular method are set as missing.

3 Data

3.1 Descriptions of Data

The data that we use for empirical analysis are drawn from a longitudinal household survey of South Korea—Korean Labor and Income Panel Study (KLIPS). This survey has been often employed to study issues of South Korea (e.g., Cho and Keum, 2004; Kang et al., 2007; Lee and Tae, 2005). KLIPS is a nationally representative longitudinal survey of Korean households that started in 1998 with 5,000 households and 13,783 individuals aged 15 or older. It is modeled after the National Longitudinal Surveys (NLS) and the Panel Study of Income Dynamics (PSID) of the U.S., and administered annually by the Korea Labor Institute (KLI), a government-sponsored research institute (Korea Labor Institute, 1998).

The KLIPS survey collects a wide range of information on families and individuals such as labor market status, earnings, family background, and demographic characteristics. From the household questionnaire of the survey, we obtain a household’s sibling composition and a child’s birth order. Starting from the third wave (survey year 2000), KLIPS collects unique information on private tutoring for children. The data contain private tutoring expenditures for each and every child in a household who attends educational institutions (including day-cares) below college. We construct our measure of educational investments for a child from this information. Specifically, we employ the monthly average expenditure on private tutoring for each individual child, which is collected in waves three to seven of KLIPS (survey years 2000 to 2004). Such a private tutoring expenditure does not include expenses on school supplies, reference books, school fees, etc. Although not verifiable using the KLIPS data, these expenses will not vary substantially by different households in the nation due to Korea’s unique system of primary and secondary education in which schooling is almost exclusively publicly provided under national education financing and even distribution of public educational resources (OECD, 1998). Variation in private tutoring expenditure can be a good proxy for that in total monetary educational investment for Korean children.

For subsequent analysis we impose several restrictions on the raw KLIPS data. First, we exclude children below elementary school age (age 7), because private tutoring expenditures for these young children may be confounded with child care expenditures. With such a restriction, we have the total number of tutoring expenditure observations for first to twelfth graders as following: 2,437 in 2000; 2,205 in 2001; 2,168 in 2002; 2,179 in 2003; 2,282 in 2004; and a total of 11,271 between 2000 and 2004. Second, those students who live in single-parent families or whose guardian is not one of the parents are also excluded, because patterns of educational investments in such families may be far from normal. However, students who cohabit with grandparents as well as both parents are included. Third, we exclude children whose mother is younger than 35 years of age, because the mother may not have completed fertility. Fourth, we focus our analysis on second-born children alone (but boys and girls, separately), who have no missing information for variables employed. The above restrictions leave us with a total of 1,942 observations for 616 second-born girls and 2,231 observations for 721 second-born boys from years 2000 to 2004. Finally, if we further exclude the observations of those whose family size is greater than three, a total of 1,893 observations are left for 603 second-born girls, and a total of 2,226 observations are left for 720 second-born boys.

3.2 Descriptive Statistics

Descriptive statistics of the main samples and their differences between daughter-first and son-first families are documented in Table 1. First three columns report statistics for second-born girls; last three columns those for second-born boys.

INSERT TABLE 1 HERE.

As for second-born girls, the monthly average expenditure on private tutoring is $W123,800$ (in 2000 constant)—approximately \$98.2. This value is fairly close to the national monthly average expenditure on tutoring ($W111,750$) reported by Ministry of Education (2000) for female students attending school from grades 1 to 12 ($W111,917$ for male students). If the sample is divided by the sex of the first-born, second-born girls in daughter-first families receive smaller educational investments than those in son-first families; the average expenditure is $W114,028$ in the daughter-first family and $W131,624$ in the son-first family. The difference is significantly different from zero. The proportion of those who receive private tutoring, having a positive

expenditure, is also lower among girls in daughter-first families (68.4 percent) than those in son-first families (72.2 percent). Such a difference is also significant. Namely, the sex of the first-born seems to affect educational investments for second-born girls. Yet it remains to be seen whether such a difference in educational investments is mediated through and caused by the difference in family size.

Family sizes are significantly different between the two types of families. As expected by son preference, the average number of children in daughter-first families is greater than that in son-first families: it is 2.48 in daughter-first families and 2.08 in son-first families.⁶ The proportion of girls who have more than one sibling is also higher in daughter-first families than in son-first families (0.483 versus 0.084). Namely, the sex of the first-born child yields a significant difference in family size for a second-born girl.

Given the differences in family size and y between daughter-first and son-first families, we can calculate a Wald estimate of the effect of increased family size on tutoring expenditures. Such an estimation supposes that the sex of the first-born child is truly exogenous to y when T is controlled for. The estimate can be calculated by $\frac{E(y_i|D_i=1)-E(y_i|D_i=0)}{E(T_i|D_i=1)-E(T_i|D_i=0)}$. The Wald estimate implies that a second-born girl receives 46.3 percent (SE 23.3) smaller educational investments if she has more than one sibling than if she has only one sibling. If we extend the analysis by controlling for other characteristics of the girl and her family, the 2SLS estimate suggests that a second-born girl receives 29.2 percent (SE 17.4) smaller educational investments if she has more than one sibling than if she has only one sibling (see Table 2). If we recall that equation (7) may hold for second-born girls and that T is positively associated with D , such an estimate is more likely to even understate negative (positive) impacts of large (small) family size than overstate them.

Concerning other variables, a student's age, grade level, parents' education level and family income are similar between daughter-first and son-first families. Parents of daughter-first families, however, are slightly older than those of son-first families. The proportion of the daughter-first family is 0.447, which fails to be significantly different from 0.5.

⁶For a combined sample of boys and girls, Lee (2007, Col (2) of Table 6) documents that the average family size of daughter-first families is greater by 0.177 than that of son-first families. If we experiment with a similar specification of Lee's using our combined sample of second-born boys and girls, we find that daughter-first families have 0.181 (SE 0.021) more children than son-first families.

As for second-born boys on the other hand, it is notable that boys in daughter-first families receive greater educational investments than those in son-first families. And it is in sharp contrast to the case of girls that family sizes fail to be significantly different between daughter-first and son-first families. Provided that the second-born child is a boy, the sex of the first-born child does not give rise to differences in the fertility decision of Korean parents; the likelihood of giving birth to more than two children in a family is quite low around an average of 6 percent regardless of the sex of the first-born. Other variables such as a student’s age, grade level, parents’ education level, age and family income are also similar between daughter-first and son-first families. Because the sex of the first-born fails to vary significantly the family size of second-born boys, Wald and 2SLS estimates cannot be convincingly produced for the boys sample; we focus on bounds analysis in order to explore the causal impact of large family size on educational investments for second-born boys.

4 Estimation Results

4.1 OLS and 2SLS Results

An an intermediate step, we estimate the following conventional model of family size effects by OLS and 2SLS:

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \epsilon_i \tag{14}$$

where y_i as a natural log of an average monthly expenditure for child i ⁷; T_i takes 0 if the total number of i ’s siblings is equal to one, and 1 if greater than one; and X_i is a vector of i ’s measured characteristics. In 2SLS, D_i is employed as an IV for T_i . Table 2 presents the estimation results for girls and boys separately.

INSERT TABLE 2 HERE.

First, family size is strongly associated with the size of educational expenditures for girls. The OLS estimate in column (1) suggests that a second-born girl receives on average a 29.5 percent smaller educational investment if she lives with more than one sibling than if she lives with

⁷In the data raw values of tutoring expenditures vary from 0 to 1,743.7. To deal with zero expenditures in the log transformation, a value of 10 is added to every child’s raw value of expenditure before taken a log. The value of 10 is chosen because it is the smallest accounting unit reported in the survey ($W10,000$) and it is about 7 to 8 percent of the mean expenditure on private tutoring. If a smaller value (e.g., 1) is added to every expenditure, however, the results are qualitatively similar.

only one sibling in the family. This amount is significant statistically as well as economically. However, it may not be a causal estimate due to endogeneity of T_i . If we rely on 2SLS so as to draw better causal estimates for the effect of large family size, the 2SLS estimate in column (3) suggests that a second-born girl receives a 29.2 percent smaller educational investment if she lives with more than one sibling than if she lives with only one sibling. As mentioned earlier, even this estimate is likely to understate negative impacts of large family size than overstate them: for second-born girls D_i is likely to be positively rather than negatively correlated with ϵ_i while being also positively correlated with T_i .⁸

The first-stage estimate in column (2) shows that sex of the first-born child is a strong predictor of family size of a second-born girl. She is more likely to live in a large family if the first-born child is a girl than if it is a boy. A second-born girl has a 41.2 percentage point higher probability of getting more than one sibling (as opposed to only one sibling) if the first-born child is a girl than if it is a boy.

Second, family size is also strongly associated with the amount of educational expenditures for boys. The OLS estimate in column (4) suggests that a second-born boy receives on average a 22 percent smaller educational investment if he lives with more than one sibling than if he lives with only one sibling. This amount is, however, marginally significant.

As observed in Table 1, sex of the first-born child fails to be a strong predictor of family size of a second-born boy. It has little correlation with the family size. Thus 2SLS methods cannot be convincingly applied to the case of boys. As expected from the extreme weakness of the IV in the first stage, the 2SLS estimate in column (6) shows an unrealistic figure of negative 11.6 (SE 35.6). Therefore, we rely on the bounds analysis in order to explore the causal impact of large family size on boys.⁹

⁸Besides a limitation from a failure to control for a child's birth order in a study of family size effects, Lee's (2007) interpretation of potential bias of his 2SLS estimates is wrong-signed. He argues that postnatal son preference yields an *overstatement* of the true causal effect of family size (p.11). However, as is clear from our setting which measures the effect on second-born boys and girls separately, using the sex of the first-born child (D_i) is more likely to *understate* the true effect of family size rather than overstate it for second-born girls. In addition, if we rule out cost-savings associated with same sex siblings, using D_i as an IV for family size is more likely to understate rather than overstate the true effect on second-born boys as well, because competition among sons within the family is likely to yield a positive (rather than negative) correlation between D_i and ϵ_i , while D_i is positively correlated with T_i .

⁹Using total spending on education as well as on private tutoring, and a combined sample of boys and girls for years 1993 to 1998, Lee's (2007) 2SLS methods find that an additional child of the two-children family decreases per-child educational investments by about 14.5 to 18.5 percent. A 2SLS estimate based on our combined sample of second-born girls and boys yields the estimated effect comparable to Lee's, although ours is more imprecisely estimated. Our 2SLS estimate suggests that an additional child decreases educational investments for a second-born child by about 10.1 percent (SE 27.1).

4.2 Results of Bounds Analysis

The estimated bounds of $E[y(t)]$ and average treatment effects (ATE) are presented in Tables 3 and 4. The estimates of the girls sample are shown in Table 3; those of the boys sample in Table 4. Figure 1 graphically displays the bounds of ATE under different combinations of the assumptions.

INSERT TABLE 3 and Figure 1 HERE.

In our specification y is a natural log of a tutoring expenditure that may vary from negative infinity to positive infinity in principle. In the data the observed y varies from 2.302 to 7.943 after a value of 10 is added to the raw value of the expenditure (see footnote 7).¹⁰ In order to make bounds analysis feasible, we arbitrarily impose lower and upper bounds of y . We set K_0 equal to 2.3, which corresponds to a zero expenditure, and K_1 equal to 10, which corresponds to a monthly expenditure of W22,016,466.¹¹

At first, if we look at the results for girls, the worst-case (WC) bounds, which are generated under no particular assumptions on statistical properties of $E[y(t)]$, suggest that ATE (i.e., $E[y(1) - y(0)]$) could be anywhere between -3.065 and 4.635. Such bounds imply that the average amount of tutoring expenditures for a second-born daughter can either decrease by as much as 307 percent or increase by 464 percent if she has more than one sibling in the family. Although such bounds are not very informative to infer the true magnitude of the effect of large family size, they are much narrower than the bounds that can range between -7.7 and 7.7 under no information whatsoever. The MTR assumption, which supposes non-positive effects of large family size, cuts the size of the WC bounds by a half, ruling out a positive ATE by design. The MTS assumption tightens the WC bounds mainly by raising the lower bound of ATE to -0.344. However, the MIV assumption by itself fails to narrow substantially the WC bounds.

If we put individual assumptions together to further tighten the bounds of ATE, the joint MIV+MTR assumption narrows the MIV bounds of ATE primarily by lowering the upper bound (UB) to -0.185; the joint MIV+MTS assumption narrows the MIV bounds by raising the lower

¹⁰Appendix Tables 1 and 2 report bounding estimates for girls and boys, respectively, when a value of 1 is added to the raw expenditure before taken a log. The results are qualitatively unaffected.

¹¹Such bounds are somewhat arbitrary. Thus we have explored how the estimated bounds vary with alternative values of K_0 and K_1 . (This sensitivity analysis is unnecessary for the MTR+MTS and MIV+MTR+MTS bounds because they are not a function of K_0 and K_1 .) Nonetheless, the primary findings of the paper are qualitatively unaffected. The results based on alternative values of K_1 are available upon request.

bound (LB) to -0.342. In addition, the joint MTR+MTS assumption considerably tightens the bounds of ATE to a range between -0.344 and 0 mainly by ruling out a positive ATE under MTS by design. Finally, the joint MIV+MTR+MTS assumption yields the tightest bounds of ATE between -0.383 and -0.144. Namely, if one accepts the three assumptions of MIV, MTR, and MTS jointly to estimate the effect of large family size, she can infer that the true size of the effect is somewhere between -0.383 and -0.144. It is reassuring that the 2SLS estimate (i.e., -0.292) in Table 2 is located within such an interval that is obtained via a different empirical approach.

The bootstrap confidence interval for the MIV+MTR+MTS bounds suggests that the UB of ATE is strictly smaller than zero at the 10 percent level of significance. Therefore, statistical evidence suggests that there exists a negative (positive) causal impact of large (small) family size on educational investments for second-born girls. If we rely on our (possibly upward-biased) 2SLS estimate as a point estimate, a 29.2 percent decrease in tutoring expenditure due to a one child increase in family size is both statistically and economically significant for second-born girls. A more conservative estimate, the UB of ATE (-0.144), also suggests a fairly strong impact of family size on educational investments for second-born girls. Further, these conclusions remain largely unaffected even if we use an alternative version of y , which is obtained by adding 1 (instead of 10) to the raw value of an expenditure before taken a log (see Appendix Table 1).

In contrast to the findings for second-born girls, the primary results for second-born boys are somewhat different, while patterns of changes in bounds with varying assumptions for boys are similar to those for girls. From Table 4, the WC bounds are not very informative in drawing the bounds of ATE, suggesting that ATE could be anywhere between -2.233 and 5.467. MTR cuts the size of the WC bounds by a half, ruling out a positive ATE. MTS tightens the WC bounds by raising the LB of ATE to -0.224. MIV under equation (7) alone does not lead to tightening the WC bounds. The joint MIV+MTR assumption tightens the MIV bounds by lowering the UB to 0; the joint MIV+MTS assumption tightens the MIV bounds by raising the LB to -0.121. The joint MTR+MTS assumption considerably tightens the bounds of ATE. Finally, the joint MIV+MTR+MTS assumption yields the tightest bounds of ATE between -0.121 and 0.

INSERT TABLE 4 HERE.

The bootstrap confidence interval for the MIV+MTR+MTS bounds suggests that the UB of ATE fails to be strictly smaller than zero at the 10 percent significance level. Although the estimated LB suggests a negative level of ATE, its 95th percentile is equal to zero; the LB also fails to be significantly different from zero. Therefore, it seems doubtful that there exist strong negative causal impacts of large family size on educational investments for boys. Even if one takes the LB of the MIV+MTR+MTS bounds of ATE seriously, the estimate suggests that second-born boys receive at most 12.1 percent smaller educational investments if they live with more than one sibling than if they live with only one sibling. To compare with the results for girls, negative (positive) impacts of large (small) family size for second-born boys seem much weaker than those for second-born girls.¹²

Such a conclusion remains largely similar even if we replace equation (7) with equation (8) for MIV. The new MIV'+MTR+MTS assumption yields the tightest bounds of ATE between -0.218 and -0.055. The bootstrap confidence interval suggests that the 95th percentiles of the LB and UB are equal to zero. Therefore, there is a doubt that family size has strong causal impacts on educational investments for second-born boys. A conservative summary of the preceding results would be that while there is a non-negligible negative (positive) effect of large (small) family size on educational investments for second-born girls, there is little if any effect of family size on educational investments for second-born boys.

4.3 Disaggregated Results of the Bounds Analysis

4.3.1 Disaggregation by the Sex of Parity 3

Our primary findings that large family size has a strong negative impact on educational investments for girls but little impact on those for boys seem to pertain to son preference traditionally shown by Korean parents. The parents treat sons relatively non-discriminately in education whether or not there exist extra siblings in the family; however, daughters suffer from the presence of extra siblings. If son preference is closely related with our empirical findings, we

¹²We extend the current bounds analysis based on second-born children alone to each of the boys and girls samples that include first-born as well as second-born children. A major disadvantage of such an extension is that MIV, which tightens bounds of ATE in combination with MTR+MTS, cannot make sense to the first-born children. Hence we calculate the WC, MTR, MTS and MTR+MTS bounds alone, failing to obtain bounds based on MIV. Not surprisingly, MTR+MTS yields the tightest bounds of ATE. For the girls sample, the MTR+MTS bounds are [-0.330, 0]; for the boys sample, they are [-0.285, 0]. Although both bounds are fairly tight and the LB of ATE for girls are larger in absolute value than that for boys, it seems difficult to draw firm conclusions about the size of the impacts of family size and the difference in impacts between girls and boys.

can expect that second-born girls and boys are exposed to different risks of large family size in education, depending on the sex of siblings who are born later. Namely, under son preference second-born girls are at a greater disadvantage in educational investments when they have a younger brother(s) than when they have a younger sister(s); however, second-born boys will be at a disadvantage that is as little—or lesser due to reduced competition with other son(s)—when they have a younger sister(s) as when they have a younger brother(s).

To explore these possibilities, we disaggregate each treatment group—a group of children with more than one sibling—of the girls and boys samples by whether the parity-3 child is a son or a daughter, keeping children with only one sibling as a control group. With such subsamples, ATE is the average effect on a second-born child of having one younger brother (or sister) relative to no younger sibling. The joint assumption bounds of ATE for the disaggregated samples are shown in Table 5 for second-born girls and Table 6 for second-born boys.

INSERT TABLES 5 and 6 HERE.

The tightest MIV+MTR+MTS bounds of ATE shown in Panel A of Table 5 suggest that a second-born girl receives at least 15.1 and at most 56.3 percent smaller educational investments if she has one younger brother than if she has no younger sibling. As with the whole girls sample, the UB of ATE is strictly smaller than zero. Thus empirical evidence supports that there is a significant negative impact of family size that is increased by a birth of a younger brother on girls' education. In contrast, the MIV+MTR+MTS bounds of ATE in Panel B seem smaller in absolute value than those in Panel A, the size of the impact being somewhere between -0.351 and -0.088. The UB of ATE, however, fails to be different from zero at the 10 percent significance level. In addition, the LB of ATE is much smaller (in absolute value) than that of Panel A. Hence, the negative impact of large family size seems to be stronger for second-born girls with a younger brother than for those with a younger sister. Nonetheless, the statistical evidence that son preference gives rise to significant variation in the impact of family size between boys and girls remains only suggestive, because there is an overlap of the two sets of MIV+MTR+MTS bounds of Panels A and B.

The estimated bounds of ATE for boys in Table 6 show similar patterns that are found for girls in Table 5, although the UBs from the boys samples fail to be significantly different from zero. The tightest MIV+MTR+MTS bounds of ATE suggest that a second-born boy receives 0

to 20.8 percent smaller educational investments if he has one younger brother than if he has no younger sibling. Although the LB of ATE implies relatively large impacts, the current bounds do not significantly rule out negligible impacts of family size on second-born boys.

If family size is increased by a birth of a sister, negative impacts of larger family size seem smaller than its impacts that result from a birth of a brother. While the UB of ATE shown in Panel B is not significantly different from zero, the LB of ATE suggests that the impact of large family size is at most a 7.7 percent decrease in educational investments if a younger sister is born to the family. Such an amount is fairly smaller in magnitude than the LB (in Panel A) of the impact of family size that is increased by a birth of a younger brother. The negative effects of large family size on second-born boys seem to be stronger if a younger brother is born to the family than if a younger sister is born. However, the evidence that son preference is strongly responsible for such findings also remain only suggestive, as the two sets of MIV+MTR+MTS bounds of Panels A and B overlap.

4.3.2 Disaggregation by Family Income

Besides the disaggregation of the analysis samples by the sex of parity 3, we divide each of the girls and boys samples by family income, using a median income (W1,978,000) as a threshold. Such a disaggregation addresses potential heterogeneity of the impacts of family size across different-income families. Because low-income families are more likely to be credit-constrained, negative impacts of large family size, if any, can be stronger in low-income families than in high-income families. Becker and Tomes (1976) predict that “observed quality and quantity would tend to be more negatively related at lower than at higher income levels (p.S147).” The joint assumption bounds of ATE for income-based subsamples are shown in Table 7 for second-born girls and in Table 8 for second-born boys.

INSERT TABLES 7 and 8 HERE.

The tightest MIV+MTR+MTS bounds of ATE shown in Panel A of Table 7 suggest that a second-born girl in a low-income family receives at least 14.1 and at most 30 percent smaller educational investments if she has more than one sibling than if she has only one sibling. As with the whole girls sample, the UB of ATE is strictly smaller than zero. Thus empirical evidence

supports that there is a strong negative impact of increased family size on education of girls in low-income families.

The MIV+MTR+MTS bounds of ATE in Panel B for girls in high-income families remain between -0.336 and -0.090. Although the LB of ATE implies relatively large impacts, the above bounds do not rule out negligible impacts of family size on girls in high-income families, because the UB of ATE fails to be strictly smaller than zero. A comparison of the MIV+MTR+MTS bounds in Panels A and B, however, does not show firm evidence that the negative effect of large family size is stronger for second-born girls in low-income families than for those in high-income families; the negative effect seems as large in low-income families as in high-income families.

The estimated bounds of ATE for boys in Table 6 show little impact of large family size on educational investments in both low-income and high-income families. As with the whole boys sample, the UB of ATE is not significantly different from zero at the 10 percent significance level. The 95th percentiles of the LBs in both Panels A and B are equal to zero. Thus there exists little causal impact of family size on educational investments for boys; this is true for boys in low-income and high-income families alike. In sum, there seems little heterogeneity of the impacts of large family size across different-income families for girls and boys alike.

5 Concluding Remarks

Employing a nonparametric bounding method and private tutoring expenditures in South Korea, this paper has examined whether large family size has a strong negative impact on educational investments for children. Our primary finding is that quantity-quality trade-offs in educational investments function in a way that varies by the sex of the child. While there is a non-negligible negative (positive) effect of large (small) family size on educational investments for second-born girls, there is little if any effect of family size on educational investments for second-born boys. Son preferences traditionally shown by Korean parents seem to underlie such empirical findings. We discover that second-born girls suffer more in education from increased family size if a younger brother is born to the family than if a younger sister is born; however, second-born boys suffer as little—or slightly less due to reduced competition among sons—when they have a younger sister as when they have a younger brother.

Several recent papers (e.g., Angrist et al., 2006; Black et al., 2005; Caceres-Delpiano, 2006; Qian, 2006) show that family size has negligible effects on the quality of children. Besides the doubts raised about exogeneity of IVs frequently employed, Liu (2007) shows that the causal relationship between quantity and quality of children depends on the measure of quality. Liu (2007) and Caceres-Delpiano (2006) empirically find the sensitivity of the estimated relationship to measures of children's quality. Although the current paper sheds light on quantity-quality trade-offs with respect to educational inputs, it does not examine a more traditional issue of whether family size has an impact on an ultimate quality such as educational outcomes of children. Such an issue may be of great interest because parents will ultimately care about the output (quality) of children rather than an input. Employing different quality measures to examine the family size effects in various dimensions would be a topic for future research.

Appendix: The Validity of the MIV Assumption for a Second-born Boy in Different Sibling Compositions

First, suppose that the effect of son preference is stronger on the distribution of educational resources among siblings than that of cost savings from same-sex siblings. Under strong son preference it is possible that parents increase total educational resources in family budget with the larger number of sons. In such a case, we suppose that they allocate the overall resources in a way that each son receives as much or more amount of resources in a family with a smaller (rather than larger) number of sons. For instance, each son in a son-son-daughter family is assumed to receive as great or greater resources than him in a son-son-son family.

Under the preceding assumptions, MIV of equation (7) holds for a second-born boy in a two-children family: he will receive as great or greater educational investments in a daughter-son family ($D_i = 1$) than in a son-son family ($D_i = 0$), whether or not parents increase overall educational budgets. In the case of three-children families, MIV of equation (7) also holds in all the four cases that follow, again irrespective of whether parents increase overall educational budgets: (1) A second-born boy is more favored in a daughter-son-daughter family ($D_i = 1$) than in a son-son-daughter family ($D_i = 0$), because he benefits from less competition among sons in the former family. (2) he is more favored in a daughter-son-son family ($D_i = 1$) than in

a son-son-son family ($D_i = 0$), and (3) in a daughter-son-daughter family ($D_i = 1$) than in a son-son-son family ($D_i = 0$), for the same reason; (4) he is equally favored in a daughter-son-son family ($D_i = 1$) and in a son-son-daughter family ($D_i = 0$), because sex composition of both families is the same..

Now, suppose that the effect of cost savings from same-sex siblings dominates that of son preference in the distribution of educational resources among siblings. Under this new assumption, MIV of equation (8) holds for a second-born boy in a two-children family: he will receive as great or greater educational investments in a son-son family ($D_i = 0$) than in a daughter-son family ($D_i = 1$) thanks to larger cost savings in the former family. In the case of three-children families, MIV of equation (8) also holds in all the four cases that follow: (1) A second-born boy is as favored or more favored in a son-son-daughter family ($D_i = 0$) than in a daughter-son-daughter family ($D_i = 1$). (2) He is more favored in a son-son-son family ($D_i = 0$) than in a daughter-son-son family ($D_i = 1$); (3) he is more favored in a son-son-son family ($D_i = 0$) than in a daughter-son-daughter family ($D_i = 1$); (4) he is equally favored in a son-son-daughter family ($D_i = 0$) and in a daughter-son-son family ($D_i = 1$).

In sum, if son preference is dominant in the distribution of family educational resources among siblings, MIV of equation (7) holds for a second-born boy both in two-children and three-children families; if cost savings from same-sex siblings are dominant, MIV of equation (8) holds both in two-children and three-children families.

References

- Almond, D., Chay, K.Y., Lee, D.S., 2005. The Costs of Low Birth Weight, *Quarterly Journal of Economics* 120 (3), 1031-1083.
- Ananat, E.O., Michaels, G., 2007. The Effect of Marital Breakup on the Income Distribution of Women with Children, CEPR Discussion Papers 6228.
- Angrist, J., Lavy, V., Schlosser, A., 2006. Multiple Experiments For The Causal Link Between The Quantity And Quality Of Children, Working Paper No.06-26, Department of Economics, Massachusetts Institute of Technology.
- Becker, G., Lewis, H.G., 1973. On the Interaction between the Quantity and Quality of Children, *Journal of Political Economy* 81 (2) part 2, S279-S288.
- Becker, G., Tomes, N., 1976. Child Endowments and the Quantity and Quality of Children, *Journal of Political Economy* 84 (4) part 2, S143-S162.
- Black, S., Devereux, P.J., Salvanes, K.G., 2005. The More the Merrier? The Effect of Family Composition on Children's Education, *Quarterly Journal of Economics* 120 (2), 669-700.
- Black, S., Devereux, P.J., Salvanes, K.G., 2007a. 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., Devereux, P.J., Salvanes, K.G., 2007b. Small Family, Smart Family? Family Size and the IQ Scores of Young Men, NBER Working Papers No.13336, National Bureau of Economic Research, Inc.
- Blake, J., 1989. *Family Size and Achievement*, Los Angeles, CA: University of California Press.
- Blundell, R., Gosling, A., Ichimura, H., Meghir, C., 2007. Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds, *Econometrica* 75 (2), 323-363.
- Caceres-Delpiano, J., 2006. The Impacts of Family Size on Investment in Child Quality, *Journal of Human Resources* 41 (4), 738-754.

- Cho, J., Keum, J., 2004. Job instability in the Korean labour market: Estimating the effects of the 1997 financial crisis, *International Labour Review* 143 (4), 373-392.
- Conley, D., Glauber, R., 2006. Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility, *Journal of Human Resources* 41 (4), 722-737.
- Dahl, G., Moretti, E., 2004. The Demand for Sons: Evidence from Divorce, Fertility, and Shotgun Marriage, NBER Working Papers No.10281, National Bureau of Economic Research, Inc.
- Das Gupta, M., Jiang, Z., Li, B., Xie, Z., Chung, W., Bae, H.O., 2003. Why is Son Preference so Persistent in East and South Asia? A Cross-country Study of China, India and the Republic of Korea, *Journal of Development Studies* 40 (2), 153-187.
- de la Croix, D., Doepke, M., 2003, Inequality and Growth: Why Differential Fertility Matters, *American Economic Review* 93 (4), 1091-1113.
- Downey, D.B., Powell, B., Steelman, L.C., Pribesh, S., 1999. Much ado about siblings: Change models, sibship size, and intellectual development, *American Sociological Review* 64 (2), 193-198.
- Galor, O., Weil, D., 2000. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond, *American Economic Review* 90 (4), 806-828.
- Gerfin, M., Schellhorn, M., 2006. Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits - Swiss evidence, *Health Economics* 15 (9), 1011-1020.
- Gonzalez, L., 2005. Nonparametric Bounds on the Returns to Language Skills, *Journal of Applied Econometrics* 20 (6), 771-795.
- Guo, G., Van Wey, L.K., 1999. Sibship Size and Intellectual Development: Is the Relationship Causal?, *American Sociological Review* 64 (2), 169-187.
- Goux, D., Maurin, E., 2005. The effect of overcrowded housing on children's performance at school, *Journal of Public Economics* 89 (5-6), 797-819.

- Hanushek, E., 1992. The Trade-off between Child Quantity and Quality, *Journal of Political Economy* 100 (1), 84-117.
- Hauser, R.M., Sewell, W.H., 1986. Sibling Size and Family Effects in Simple Models of Education, Occupational Status, *Journal of Labor Economics* 4 (3), S83-115
- Hotz, V.J., Mullin, C.H., Sanders, S.G., 1997. Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analysing the Effects of Teenage Childbearing, *Review of Economic Studies* 64 (4), 575-603.
- Kang, C., 2007. Does Money Matter? The Effect of Private Tutoring Expenditures on Academic Performance of South Korean Students, Working Paper, Department of Economics, National University of Singapore.
- Kang, C., Park, C., Lee, M.J., 2007. Effects of ability mixing in high school on adulthood earnings: Quasiexperimental evidence from South Korea, *Journal of Population Economics* 20 (2), 269-297
- Kim, S., Lee, J.-H., 2001. Demand for education and developmental state: Private tutoring in South Korea, Social Science Research Network Electronic Paper Collection: <http://ssrn.com/abstract=268284>.
- Kreider, B., Pepper, J.V., 2007. Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors, *Journal of the American Statistical Association* 102 (478), 432-441.
- Korea Labor Institute, 1998. Reports on the Korean Labor and Income Panel Study (various years), Korea Labor Institute, Seoul: Korea.
- Lechner, J., 1999. Nonparametric bounds on employment and income effects of continuous vocational training in East Germany, *Econometrics Journal* 2 (1), 1-28.
- Lee, J., 2007. Sibling Size and Investment in Children's Education: An Asian Instrument, *Journal of Population Economics*, forthcoming.
- Lee, M.J., Tae, Y.H., 2005. Analysis of labour participation behaviour of Korean women with dynamic probit and conditional logit, *Oxford Bulletin of Economics and Statistics* 67 (1), 71-91.

- Li, H., Zhang, J., 2007. Do High Birth Rates Hamper Economic Growth, *Review of Economics and Statistics* 89 (1), 110-117.
- Liu, H., 2007. Is There a Quality-Quantity Tradeoff? Evidence from the Relaxation of China's One Child Policy, Mimeo, Department of Economics, National University of Singapore.
- Manski, C.F., 1990. Nonparametric Bounds on Treatment Effects, *American Economic Review* 80 (2), 319-323.
- Manski, C.F., 1997. Monotone Treatment Response, *Econometrica* 65(6), 1311-1334.
- Manski, C.F., Nagin, D.S., 1998. Bounding Disagreements About Treatment Effects: A Case Study of Sentencing and Recidivism, *Sociological Methodology* 28, 99-137
- Manski, C.F., Pepper, J.V., 2000. Monotone Instrumental Variables: With an Application to the Returns to Schooling, *Econometrica* 68 (4), 997-1012.
- Ministry of Education, 2000. Sakyoyuk Siltae Josa [Survey on Private Tutoring in 2000], Seoul, South Korea (in Korean).
- Moav, O., 2005. Cheap Children and the Persistence of Poverty, *Economic Journal* 115 (500), 88-110.
- Moffitt, R., 2005. Remarks on the Analysis of Causal Relationships in Population Research, *Demography* 42 (1), 91-108.
- OECD (Organization for Economic Co-Operation and Development), 1998. *Reviews of National Policies for Education: Korea*, OECD Publications, Paris.
- Park, C.B., Cho, N.H., Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea, *Population and Development Review* 21 (1), 59-84.
- Pepper, J.V., 2000. The Intergenerational Transmission Of Welfare Receipt: A Nonparametric Bounds Analysis, *Review of Economics and Statistics* 82 (3), 472-488.
- Qian, N., 2006. Quantity-Quality and the One Child Policy: The Positive Effect of Family Size on School Enrollment in China, Working Paper, Department of Economics, Brown University.

Rosenzweig, M., Wolpin, K., 1980. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment, *Econometrica* 48 (1), 227-240.

Rosenzweig, M., Wolpin, K., 2000. Natural 'Natural Experiments' in Economics, *Journal of Economic Literature* 38 (4), 827-874.

Rosenzweig, M., Zhang, J., 2006. Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight, and China's 'One Child' Policy, IZA Discussion Paper, No. 2082.

Figure 1: Bounds of the Average Treatment Effect

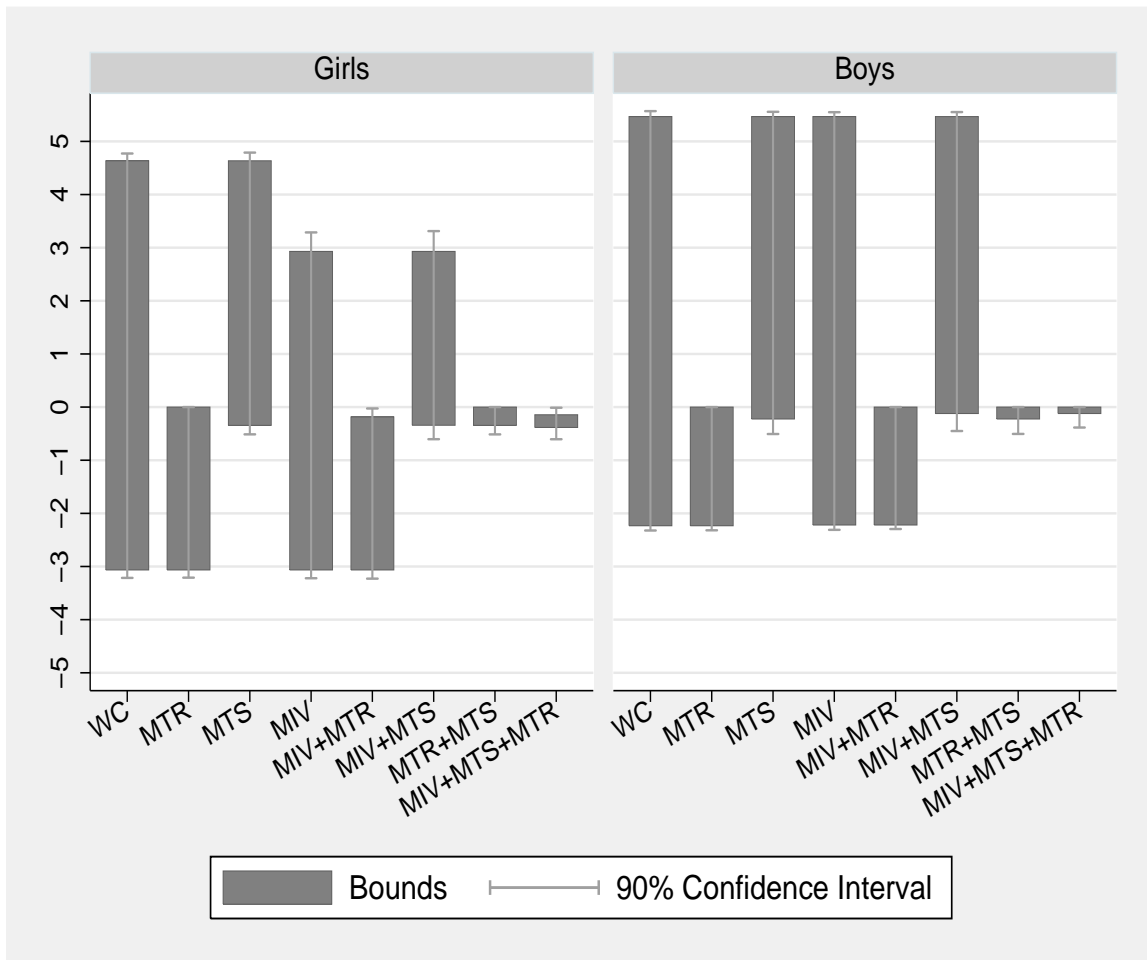


Table 1: Descriptive Statistics of the Main Samples

Variables	Girls Sample (N=1,893)			Boys Sample (N=2,226)		
	Total Sample	Daughter-first family	Differences with son-first family	Total Sample	Daughter-first family	Differences with son-first family
	Mean (S.D.)	Mean (S.D.)	Mean (S.E.)	Mean (S.D.)	Mean (S.D.)	Mean (S.E.)
Spending on tutoring (W1,000)	123.8 (151.5)	114.0 (153.6)	-17.596 (6.993)**	135.4 (174.8)	145.4 (205.4)	20.725 (7.405)**
Log(spending on tutoring)	4.217 (1.343)	4.115 (1.339)	-0.185 (0.062)**	4.286 (1.358)	4.338 (1.360)	0.107 (0.058)*
Any tutoring (Yes=1)	0.705 (0.456)	0.684 (0.465)	-0.038 (0.021)*	0.720 (0.449)	0.734 (0.442)	0.030 (0.019)
Family size	2.263 (0.440)	2.483 (0.500)	0.399 (0.018)**	2.059 (0.236)	2.055 (0.229)	-0.008 (0.010)
More than one sibling (Yes=1)	0.263 (0.440)	0.483 (0.500)	0.399 (0.018)**	0.059 (0.236)	0.055 (0.229)	-0.008 (0.010)
Age	12.485 (3.346)	12.496 (3.355)	0.019 (0.155)	12.424 (3.298)	12.423 (3.256)	-0.001 (0.140)
Grade	6.598 (3.383)	6.617 (3.383)	0.035 (0.156)	6.486 (3.325)	6.499 (3.270)	0.027 (0.141)
Father's education	12.311 (2.762)	12.306 (2.751)	-0.009 (0.128)	12.512 (2.975)	12.447 (3.034)	-0.135 (0.126)
Mother's education	11.288 (2.463)	11.320 (2.525)	0.057 (0.114)	11.706 (2.606)	11.797 (2.541)	0.190 (0.110)*
Father's age	43.795 (3.975)	44.141 (4.098)	0.626 (0.183)**	43.748 (4.163)	43.757 (4.113)	0.019 (0.177)
Mother's age	40.640 (3.674)	40.848 (3.877)	0.376 (0.170)**	40.674 (3.686)	40.792 (3.659)	0.246 (0.156)
Family income (W1,000)	2,145 (1,351)	2,134 (1,441)	-20.37 (62.46)	2,222 (1,483)	2,223 (1,488)	1.641 (62.94)
Log(family income)	7.528 (0.695)	7.483 (0.780)	-0.082 (0.032)**	7.531 (0.757)	7.530 (0.760)	-0.004 (0.032)
Live with grandparents (Yes=1)	0.024 (0.152)	0.017 (0.128)	-0.013 (0.007)*	0.016 (0.124)	0.013 (0.113)	-0.006 (0.005)
Daughter-first family (Yes=1)	0.447 (0.497)			0.519 (0.500)		

Note: * significant at the 10% level; ** significant at the 5% level

Table 2: OLS and 2SLS estimates of the Average Treatment Effect

Estimation method: Dependent variable:	Girls Sample (N=1,893)			Boys Sample (N=2,226)		
	OLS		2SLS	OLS		2SLS
	Log(Tutoring Expenditure) (1)	More than one sibling (Yes=1) (2)	Log(Tutoring Expenditure) (3)	Log(Tutoring Expenditure) (4)	More than one sibling (Yes=1) (5)	Log(Tutoring Expenditure) (6)
More than one sibling (Yes=1)	-0.295 (0.084)**		-0.292 (0.174)*	-0.220 (0.130)*		-11.60 (35.60)
Daughter-first (Yes=1)		0.412 (0.036)**			-0.006 (0.019)	
Father's education	0.047 (0.018)**	-0.008 (0.009)	0.047 (0.018)**	0.070 (0.017)**	0.001 (0.006)	0.087 (0.086)
Mother's education	0.083 (0.021)**	0.015 (0.010)*	0.083 (0.021)**	0.071 (0.021)**	0.001 (0.006)	0.077 (0.077)
Father's age	-0.014 (0.016)	-0.026 (0.007)**	-0.014 (0.016)	-0.001 (0.016)	0.000 (0.004)	0.000 (0.045)
Mother's age	-0.008 (0.016)	-0.007 (0.007)	-0.008 (0.016)	0.011 (0.016)	-0.006 (0.004)	-0.053 (0.215)
Log(family income)	0.504 (0.048)**	-0.027 (0.019)	0.504 (0.049)**	0.425 (0.048)**	-0.020 (0.013)	0.202 (0.712)
Grandparents (Yes=1)	-0.216 (0.225)	0.103 (0.074)	-0.216 (0.224)	-0.090 (0.265)	0.046 (0.067)	0.436 (1.784)
Intercept	-0.005 (0.644)	1.478 (0.277)**	-0.009 (0.688)	-1.131 (0.666)*	0.371 (0.135)**	3.069 (13.14)
F(excluded IV)		130.22			0.11	
R-square	0.268	0.258	0.268	0.023	0.290	-

Note: * significant at the 10% level; ** significant at the 5% level

Each regression controls for dummies for a student's grade, residential region and the wave of KLIPS in addition to the variables reported.

Table 3: Bounds of $E[y(t)]$ and the Average Treatment Effect for Girls ($N = 1,893$)

Assumptions:	Worst-case (WC)				Monotone treatment response (MTR)			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.780	5.802	3.691	5.995	4.217	5.802	4.136	5.989
$E[y(1)]$	2.737	8.415	2.673	8.618	2.737	4.217	2.669	4.297
$E[y(1) - y(0)]$	-3.065	4.635	-3.215	4.772	-3.065	0.000	-3.210	0.000
Assumptions:	Monotone treatment selection (MTS)				Monotone IV (MIV)			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.780	4.307	3.683	4.390	4.176	5.802	4.052	6.009
$E[y(1)]$	3.963	8.415	3.804	8.603	2.737	7.105	2.665	7.437
$E[y(1) - y(0)]$	-0.344	4.635	-0.513	4.789	-3.065	2.929	-3.221	3.287
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.299	5.802	4.198	6.003	4.176	4.217	4.043	4.330
$E[y(1)]$	2.737	4.115	2.671	4.222	3.875	7.105	3.646	7.473
$E[y(1) - y(0)]$	-3.065	-0.185	-3.229	-0.026	-0.342	2.929	-0.604	3.312
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.217	4.307	4.136	4.396	4.299	4.217	4.190	4.356
$E[y(1)]$	3.963	4.217	3.817	4.291	3.875	4.115	3.673	4.229
$E[y(1) - y(0)]$	-0.344	0.000	-0.514	0.000	-	-	-	-
Alternative					-0.383	-0.144	-0.605	-0.012
$E[y(1) - y(0)]$								

Table 4: Bounds of $E[y(t)]$ and the Average Treatment Effect for Boys ($N = 2, 226$)

Assumptions:	Worst-case (WC)				Monotone treatment response (MTR)			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.181	4.638	4.100	4.748	4.286	4.638	4.216	4.748
$E[y(1)]$	2.405	9.649	2.370	9.742	2.405	4.286	2.374	4.352
$E[y(1) - y(0)]$	-2.233	5.467	-2.323	5.570	-2.233	0.000	-2.319	0.000
Assumptions:	Monotone treatment selection (MTS)				MTR+MTS			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.181	4.300	4.100	4.370	4.286	4.300	4.209	4.371
$E[y(1)]$	4.075	9.649	3.797	9.733	4.075	4.286	3.816	4.356
$E[y(1) - y(0)]$	-0.224	5.467	-0.507	5.557	-0.224	0.000	-0.505	0.000
Assumptions:	Monotone IV (MIV) under (7)				MIV' under equation (8)			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.181	4.638	4.100	4.749	4.246	4.600	4.132	4.725
$E[y(1)]$	2.419	9.649	2.380	9.726	2.405	9.630	2.373	9.725
$E[y(1) - y(0)]$	-2.219	5.467	-2.310	5.550	-2.195	5.384	-2.299	5.512
Assumptions:	MIV+MTR				MIV' +MTR			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.286	4.638	4.219	4.738	4.338	4.600	4.244	4.723
$E[y(1)]$	2.419	4.286	2.381	4.356	2.405	4.231	2.374	4.322
$E[y(1) - y(0)]$	-2.219	0.000	-2.295	0.000	-2.195	-0.107	-2.293	0.000
Assumptions:	MIV+MTS				MIV' +MTS			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.181	4.300	4.100	4.374	4.246	4.234	4.139	4.327
$E[y(1)]$	4.179	9.649	3.864	9.731	4.068	9.630	3.778	9.715
$E[y(1) - y(0)]$	-0.121	5.467	-0.450	5.552	-	-	-	-
Alternative					-0.172	5.390	-0.471	5.514
$E[y(1) - y(0)]$								
Assumptions:	MIV+MTR+MTS				MIV' +MTR+MTS			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	4.286	4.300	4.215	4.367	4.338	4.234	4.250	4.339
$E[y(1)]$	4.179	4.286	3.927	4.355	4.068	4.231	3.792	4.333
$E[y(1) - y(0)]$	-0.121	0.000	-0.385	0.000	-	-	-	-
Alternative					-0.218	-0.055	-0.500	0.000
$E[y(1) - y(0)]$								

Table 5: Bounds for the Girls Sample Disaggregated by the Sex of Parity 3

Panel A: Son in Parity 3 ($N = 1,636$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.336	5.142	4.235	5.320	4.316	4.217	4.211	4.352
$E[y(1)]$	2.538	4.125	2.483	4.247	3.713	7.946	3.306	8.281
$E[y(1) - y(0)]$	-2.604	-0.210	-2.748	-0.043	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.553	3.679	-1.038	4.038
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.250	4.307	4.170	4.403	4.336	4.217	4.223	4.348
$E[y(1)]$	3.922	4.250	3.716	4.330	3.713	4.125	3.271	4.243
$E[y(1) - y(0)]$	-0.385	0.000	-0.614	0.000	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.563	-0.151	-1.004	-0.012
Panel B: Daughter in Parity 3 ($N = 1,653$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.310	5.192	4.212	5.371	4.203	4.217	4.084	4.349
$E[y(1)]$	2.565	4.175	2.513	4.287	3.913	8.248	3.633	8.593
$E[y(1) - y(0)]$	-2.627	-0.135	-2.771	0.000	-0.304	4.045	-0.614	4.407
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.259	4.307	4.179	4.399	4.310	4.217	4.212	4.345
$E[y(1)]$	4.002	4.259	3.795	4.346	3.913	4.175	3.639	4.292
$E[y(1) - y(0)]$	-0.305	0.000	-0.538	0.000	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.351	-0.088	-0.626	0.000

Table 6: Bounds for the Boys Sample Disaggregated by the Sex of Parity 3

Panel A: Son in Parity 3 ($N = 2,132$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctile	95 pctile			5 pctile	95 pctile
$E[y(0)]$	4.293	4.401	4.215	4.477	4.264	4.300	4.191	4.375
$E[y(1)]$	2.340	4.293	2.314	4.361	4.092	9.891	3.425	9.935
$E[y(1) - y(0)]$	-2.061	0.000	-2.132	0.000	-0.208	5.627	-0.861	5.699
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctile	95 pctile			5 pctile	95 pctile
$E[y(0)]$	4.293	4.300	4.217	4.374	4.293	4.300	4.226	4.368
$E[y(1)]$	3.899	4.293	3.385	4.368	4.092	4.293	3.460	4.361
$E[y(1) - y(0)]$	-0.401	0.000	-0.907	0.000	-0.208	0.000	-0.841	0.000
Panel B: Daughter in Parity 3 ($N = 2,188$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctile	95 pctile			5 pctile	95 pctile
$E[y(0)]$	4.293	4.545	4.227	4.642	4.214	4.300	4.140	4.377
$E[y(1)]$	2.383	4.293	2.357	4.357	4.223	9.745	3.941	9.822
$E[y(1) - y(0)]$	-2.162	0.000	-2.240	0.000	-0.077	5.532	-0.353	5.611
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctile	95 pctile			5 pctile	95 pctile
$E[y(0)]$	4.293	4.300	4.216	4.378	4.293	4.300	4.220	4.375
$E[y(1)]$	4.147	4.293	3.847	4.370	4.223	4.293	3.940	4.364
$E[y(1) - y(0)]$	-0.153	0.000	-0.466	0.000	-0.077	0.000	-0.378	0.000

Table 7: Bounds for the Girls Sample Disaggregated by Family Income

Panel A: Family Income Below Median ($N = 947$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.888	5.645	3.741	5.924	3.784	3.686	3.644	3.847
$E[y(1)]$	2.669	3.646	2.595	3.763	3.487	6.665	3.211	7.152
$E[y(1) - y(0)]$	-2.976	-0.242	-3.220	-0.057	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.248	2.930	-0.549	3.385
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.777	3.862	3.674	3.974	3.888	3.686	3.754	3.868
$E[y(1)]$	3.570	3.777	3.388	3.876	3.487	3.646	3.218	3.768
$E[y(1) - y(0)]$	-0.293	0.000	-0.515	0.000	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.300	-0.141	-0.569	-0.009
Panel B: Family Income Above Median ($N = 946$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.697	5.959	4.601	6.173	4.554	4.693	4.414	4.798
$E[y(1)]$	2.805	4.604	2.708	4.708	4.359	7.566	4.149	7.923
$E[y(1) - y(0)]$	-3.154	-0.092	-3.305	0.000	-0.334	3.012	-0.569	3.417
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	4.656	4.719	4.578	4.821	4.697	4.693	4.600	4.804
$E[y(1)]$	4.451	4.656	4.289	4.740	4.359	4.604	4.153	4.713
$E[y(1) - y(0)]$	-0.268	0.000	-0.468	0.000	-	-	-	-
Alternative $E[y(1) - y(0)]$					-0.336	-0.090	-0.573	0.000

Table 8: Bounds for the Boys Sample Disaggregated by Family Income

Panel A: Family Income Below Median ($N = 1,073$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctl	95 pctl			5 pctl	95 pctl
$E[y(0)]$	3.798	4.206	3.714	4.352	3.710	3.807	3.610	3.907
$E[y(1)]$	2.402	3.798	2.363	3.883	3.875	9.581	3.438	9.699
$E[y(1) - y(0)]$	-1.804	0.000	-1.931	0.000	0.068	5.871	-0.399	6.001
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctl	95 pctl			5 pctl	95 pctl
$E[y(0)]$	3.798	3.807	3.696	3.902	3.798	3.807	3.709	3.898
$E[y(1)]$	3.664	3.798	3.382	3.886	3.875	3.798	3.467	3.883
$E[y(1) - y(0)]$	-0.143	0.000	-0.463	0.000	0.000	0.000	-0.359	0.000
Panel B: Family Income Above Median ($N = 1,153$)								
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctl	95 pctl			5 pctl	95 pctl
$E[y(0)]$	4.741	5.040	4.660	5.143	4.619	4.753	4.535	4.834
$E[y(1)]$	2.436	4.741	2.385	4.831	4.535	9.701	4.239	9.802
$E[y(1) - y(0)]$	-2.605	0.000	-2.694	0.000	-0.218	5.082	-0.540	5.179
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB	UB	LB	UB	LB	UB
			5 pctl	95 pctl			5 pctl	95 pctl
$E[y(0)]$	4.741	4.753	4.661	4.843	4.741	4.753	4.656	4.829
$E[y(1)]$	4.526	4.741	4.105	4.823	4.535	4.741	4.176	4.815
$E[y(1) - y(0)]$	-0.228	0.000	-0.642	0.000	-0.218	0.000	-0.596	0.000

Appendix Table 1: Bounds of $E[y(t)]$ and the Average Treatment Effect for Girls with an Alternative y

Assumptions:	Worst-case (WC)				Monotone treatment response (MTR)			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	2.671	5.297	2.514	5.528	3.483	5.297	3.349	5.523
$E[y(1)]$	0.812	8.187	0.694	8.419	0.812	3.483	0.691	3.623
$E[y(1) - y(0)]$	-4.484	5.516	-4.651	5.671	-4.484	0.000	-4.641	0.000
Assumptions:	Monotone treatment selection (MTS)				Monotone IV (MIV)			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	2.671	3.622	2.496	3.766	3.374	5.297	3.149	5.541
$E[y(1)]$	3.094	8.187	2.818	8.411	0.812	6.699	0.684	7.086
$E[y(1) - y(0)]$	-0.528	5.516	-0.822	5.688	-4.484	3.325	-4.655	3.778
Assumptions:	MIV+MTR				MIV+MTS			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	3.607	5.297	3.433	5.537	3.374	3.486	3.145	3.681
$E[y(1)]$	0.812	3.330	0.690	3.512	2.946	6.699	2.531	7.122
$E[y(1) - y(0)]$	-4.484	-0.277	-4.664	-0.002	-0.540	3.325	-1.003	3.799
Assumptions:	MTR+MTS				MIV+MTR+MTS			
	LB	UB	LB 5 pctile	UB 95 pctile	LB	UB	LB 5 pctile	UB 95 pctile
$E[y(0)]$	3.483	3.622	3.344	3.781	3.607	3.486	3.425	3.719
$E[y(1)]$	3.094	3.483	2.831	3.617	2.946	3.330	2.580	3.515
$E[y(1) - y(0)]$	-0.528	0.000	-0.822	0.000	-	-	-	-
Alternative					-0.600	-0.216	-0.995	0.000
$E[y(1) - y(0)]$								

Appendix Table 2: Bounds of $E[y(t)]$ and the Average Treatment Effect for Boys with an Alternative y

Assumptions:	Worst-case (WC)				Monotone treatment response (MTR)			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.391	3.984	3.257	4.129	3.587	3.984	3.462	4.132
$E[y(1)]$	0.196	9.603	0.131	9.707	0.196	3.587	0.139	3.703
$E[y(1) - y(0)]$	-3.789	6.211	-3.913	6.354	-3.789	0.000	-3.905	0.000
Assumptions:	Monotone treatment selection (MTS)				MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.391	3.605	3.257	3.731	3.587	3.605	3.458	3.718
$E[y(1)]$	3.299	9.603	2.840	9.697	3.299	3.587	2.818	3.701
$E[y(1) - y(0)]$	-0.306	6.211	-0.788	6.338	-0.306	0.000	-0.781	0.000
Assumptions:	Monotone IV (MIV) under (7)				MIV' under equation (8)			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.391	3.984	3.255	4.138	3.497	3.913	3.311	4.088
$E[y(1)]$	0.219	9.603	0.148	9.691	0.196	9.584	0.134	9.692
$E[y(1) - y(0)]$	-3.766	6.211	-3.887	6.323	-3.717	6.087	-3.870	6.270
Assumptions:	MIV+MTR				MIV' +MTR			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.587	3.984	3.473	4.123	3.671	3.913	3.519	4.091
$E[y(1)]$	0.219	3.587	0.152	3.703	0.196	3.496	0.139	3.653
$E[y(1) - y(0)]$	-3.766	0.000	-3.877	0.000	-3.717	-0.174	-3.868	0.000
Assumptions:	MIV+MTS				MIV' +MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.391	3.605	3.251	3.724	3.497	3.500	3.326	3.657
$E[y(1)]$	3.445	9.603	2.913	9.695	3.289	9.584	2.814	9.680
$E[y(1) - y(0)]$	-0.160	6.211	-0.706	6.330	-0.211	6.087	-0.664	6.248
Assumptions:	MIV+MTR+MTS				MIV' +MTR+MTS			
	LB	UB	LB 5 pctl	UB 95 pctl	LB	UB	LB 5 pctl	UB 95 pctl
$E[y(0)]$	3.587	3.605	3.461	3.723	3.671	3.500	3.522	3.667
$E[y(1)]$	3.445	3.587	3.016	3.699	3.289	3.496	2.885	3.644
$E[y(1) - y(0)]$	-0.160	0.000	-0.594	0.000	-	-	-	-
Alternative					-0.296	-0.089	-0.746	0.000
$E[y(1) - y(0)]$								