



**The Quantity and Quality of Children:  
A Semi-Parametric Bayesian IV Approach**

by

Sylvia FRÜHWIRTH-SCHNATTER  
Martin HALLA\*)  
Alexandra POSEKANY  
Gerald J. PRUCKNER  
Thomas SCHÖBER

Working Paper No. 1402

March 2014

Supported by the  
Austrian Science Funds

**FWF**

---

**The Austrian Center for Labor  
Economics and the Analysis of  
the Welfare State**

JKU Linz  
Department of Economics  
Altenberger Strasse 69  
4040 Linz, Austria  
[www.laborrn.at](http://www.laborrn.at)

Corresponding author: [martin.halla@jku.at](mailto:martin.halla@jku.at)  
Phone: +43 (0)732 2468 8706

# The Quantity and Quality of Children: A Semi-Parametric Bayesian IV Approach

SYLVIA FRÜHWIRTH-SCHNATTER<sup>†</sup>

MARTIN HALLA<sup>‡§</sup>

ALEXANDRA POSEKANY<sup>†</sup>

GERALD J. PRUCKNER<sup>‡</sup>

THOMAS SCHOBER<sup>‡</sup>

March 2, 2014<sup>¶</sup>

## Abstract

Prior empirical research on the theoretically proposed interaction between the quantity and the quality of children builds on exogenous variation in family size due to twin births and focuses on human capital outcomes. The typical finding can be described as a statistically nonsignificant two-stage least squares (2SLS) estimate, with substantial standard errors. We regard these conclusions of no empirical support for the quantity-quality trade-off as premature and, therefore, extend the empirical approach in two ways. First, we add health as an additional outcome dimension. Second, we apply a semi-parametric Bayesian IV approach for econometric inference. Our estimation results substantiate the finding of a zero effect: we provide estimates with an increased precision by a factor of approximately twenty-three, for a broader set of outcomes.

*JEL Classification:* J13, C26, C11, I20, J20, I10.

*Keywords:* Quantity-quality model of fertility, family size, human capital, health, semi-parametric Bayesian IV approach.

---

<sup>†</sup>Vienna University of Economics and Business, Institute for Statistics and Mathematics, Austria.

<sup>‡</sup>Johannes Kepler University of Linz, Department of Economics, Austria.

<sup>§</sup>Institute for the Study of Labor (IZA), Bonn, Germany.

<sup>¶</sup>Corresponding author: Martin Halla, Johannes Kepler University of Linz, Department of Economics, Altenberger Strasse 69, 4040 Linz, Austria; ph.: +43 (0)732 2468 8706; email: martin.halla@jku.at. This research was funded by the Austrian Science Fund (FWF): National Research Network S103, The Austrian Center for Labor Economics and the Analysis of the Welfare State.

# 1 Introduction

Motivated by the theoretically proposed interaction between the quantity and quality of children (Becker and Lewis, 1973; Willis, 1973; Becker and Tomes, 1976), economic scholars have an ongoing interest in the empirical effect of family size on child outcomes. While naïve estimations consistently show that children’s human capital outcomes are negatively affected by a greater family size, more recent instrumental variable (IV) approaches rarely find statistically significant point estimates and tend to reject the quantity-quality trade-off.<sup>1</sup>

We argue that the conclusions drawn from these recent IV approaches may be premature. First, the empirical studies have primarily examined educational and labor market outcomes. Although we agree that these two dimensions are important, the quantity-quality trade-off may still materialize in other domains. For example, we consider health to be an important dimension.<sup>2</sup>

Second, it is debatable whether existing IV estimates on human capital outcomes are sufficiently precise to allow a conclusion of zero effects. The standard approach is based on two-stage least squares (2SLS) estimates that exploit exogenous variations in family size due to twin births and preferences for a mixed sibling-sex composition. A key concern with 2SLS, even with very large sample sizes, is whether the estimates are precise enough to be informative. For instance, Angrist *et al.* (2010) conducted a study based on Israeli data using the twin instrument and found an estimated effect of family size on college attendance of plus 1.7 percentage points, with a standard error of 5.2. This effect is clearly statistically nonsignificant. However, the corresponding 95 percent confidence interval is large, ranging from minus 8.5 to plus 11.9. We do not believe this estimate justifies the strong conclusion that family size has no impact on child outcomes.<sup>3</sup>

This study contributes to the existing body of literature along two dimensions. First, we use administrative data sources from Austria to extend the conventional human capital

---

<sup>1</sup>The evidence from developed countries comprises educational outcomes and IQ scores in Norway (Black *et al.*, 2005, 2010), educational outcomes in the USA (Cáceres-Delpiano, 2006), educational, labor market, and family outcomes in Israel (Angrist *et al.*, 2010), and educational and labor market outcomes in Sweden (Åslund and Grönqvist, 2010). The estimation results from these studies are summarized in Table 1. For a discussion of estimates from low- and medium-income countries see, for instance, Quian (2013).

<sup>2</sup>We are only aware of two design-based papers (twin instrument) based on data from developed countries. Hatton and Martin (2010) find a negative effect on the height of British children born in the interwar period, but no effect on their body mass index. Then, Lundborg *et al.* (2013) identify a positive effect on both height and a global health measure using Swedish data that covers more recent birth cohorts. Details are provided in Table 1.

<sup>3</sup>Most recently, Bagger *et al.* (2013) presented a theoretical extension of the quantity-quality model that allows for the differential treatment of children by birth order. In their empirical test, using the twin instrument in a 2SLS framework, they find a statistically significant negative effect of family size on children’s *average* educational outcomes. In other studies, Mogstad and Wiswall (2012) and Brinch *et al.* (2012) discuss the functional form assumption employed by previous empirical studies and provide evidence of non-linear effects. However, this aspect is less relevant in our context as we use data from a low-fertility environment.

outcomes to include health outcomes. In particular, we complement college attendance, employment, occupation, and wages with health-related outcome variables such as the number of days of hospitalization and expenditure on medication and outpatient medical attendance. Second, to tackle the problem of the large standard errors of the 2SLS approach, we apply the semi-parametric Bayesian IV approach advocated by Conley *et al.* (2008). The results of this estimation method are substantially more precise than 2SLS results. To demonstrate this, for each outcome, we compare the relative performance of the conventional IV approach (2SLS) to that of the standard and semi-parametric Bayesian IV approaches.

*Theoretical considerations* The quantity-quality model of children assumes egalitarian parents who aim to invest in all their children equally. The interaction of child quality and child quantity in the budget constraint means that an exogenous increase in the quantity of children increases the marginal cost of investing in their quality. Consequently, children from larger families are predicted to receive less parental investment and to have worse outcomes. In the case of human capital outcomes, there is an obvious causal chain running from pecuniary and non-pecuniary investment, via educational outcomes, to labor market performance. In terms of health outcomes, multiple channels are possible. First, larger families may not be able to afford the same quality of food and healthcare services. Equivalently, it may be difficult for them to find the time for preventive and/or curative doctor visits. Alternatively, time constraints could lead to deficiencies in children's health education. Thus, children from larger families may be more prone to an unhealthy lifestyle than their counterparts from smaller families.

Certainly, the simple and elegant quantity-quality model of children exposes an important mechanism, which applies in either outcome domain. However, it disregards other potentially important channels. In particular, an additional sibling might affect a firstborn beyond the negative reinforcing mechanism between quantity and quality. Put differently, an additional sibling may not only be a rival, but also an important resource. Developmental psychology stresses that caring for younger siblings may yield some developmental benefits for older siblings. Family psychology literature discusses how siblings can provide both financial and emotional support (Pollet and Hoben, 2011). Concerning health, the social interaction channel is a well-known subject in medical literature. The so-called hygiene hypothesis states that a lack of early childhood exposure to infectious agents, particular microorganisms and parasites, may increase a person's susceptibility to allergic disorders and autoimmune diseases by suppressing the natural development of the immune system (see, for instance, Strachan, 1989, 2000). Since a larger family size implies a greater exposure to pathogens, parasites, bacteria, and viruses, the hygiene hypothesis predicts that family size has a positive influence on the health of children. However, medical and epidemiological literature on this subject is inconclusive, not least

because it disregards the endogeneity of family size.<sup>4</sup>

*Bayesian inference* Unlike the conventional approach, the Bayesian paradigm assumes distributions for observed variables as well as parameters. In particular, a Bayesian inference requires formulating the conditional distribution of the data, given model parameters (i.e., the likelihood function), and a prior distribution for the model parameters themselves. For Bayesian IV models, two different approaches have been suggested to derive the conditional distribution of the data from the joint distribution of the outcome and endogenous variables. In the standard Bayesian IV estimation, the likelihood function is derived by assuming a bivariate normal distribution for the outcome and endogenous variables, as described in Rossi *et al.* (2005). In contrast, the semi-parametric Bayesian approaches aim to avoid rigid choices of data distributions by allowing a wider class of possible likelihood functions. This makes them well-suited and flexible alternatives for data that violate the normal distribution assumptions for outcome and endogenous variables.<sup>5</sup>

The idea behind the semi-parametric approach is that any distribution can be represented by mixing a sufficient number of normal distributions. If only a single normal distribution is necessary, then the standard Bayesian approach is the optimal choice. Mixing several normal distribution components can represent more complex scenarios. In other words, the less similar a distribution is to the normal distribution, the greater the number of mixture components it requires. In such a non-normal case, we cannot rely on results from methods that assume normality, such as the standard Bayesian IV or the typical confidence interval estimation of conventional IV. Our human capital outcomes and, in particular, our health outcomes represent complex distributions that require many components. Therefore, we have more confidence in the findings of a semi-parametric Bayesian approach than we do in those of the standard Bayesian or conventional IV approaches.

Conley *et al.* (2008) compare their semi-parametric Bayesian IV approach to the standard Bayesian and conventional IV methods and find it to be more efficient for non-normal errors. In particular, they compare the semi-parametric Bayesian IV approach to point

---

<sup>4</sup>A series of medical studies in several European countries report a negative association between family size and allergic diseases. These studies postulate that family size is a protective factor for allergies (Von Mutius *et al.*, 1994; Jarvis *et al.*, 1997; Strachan *et al.*, 1997; Forastiere *et al.*, 1997; Svanes *et al.*, 1999). Wickens *et al.* (1999) take a more critical view of the impact of family size on the increase in the prevalence of asthma and hay fever. The authors show that the relative change in family size in England/Wales and in New Zealand between 1961 and 1991 did not appear to explain much of the increase in asthma and hay fever prevalence. In reviewing 53 different studies, Karmaus and Botezan (2002) support the “protective effect” of a greater number of siblings on the risk of atopic eczema, asthma and wheezing, hay fever, and allergic sensitization. However, the authors point out that this result is based exclusively on data provided by epidemiological associations and that the question of causal factors has not yet been answered.

<sup>5</sup>Books by Zellner (1971) and Rossi *et al.* (2005) review and present Bayesian statistical inference in econometrics and marketing, including the Bayesian IV approach. Kleibergen and Zivot (2003) compare the Bayesian and conventional approaches to inferring IV models. Hoogerheide *et al.* (2007), Lopes and Polson (2014), and Zellner *et al.* (2012) discuss using Bayesian models for the IV problem and choosing the prior distributions for the parameters. Chamberlain and Imbens (2003) and Conley *et al.* (2008) advocate semi-parametric approaches to Bayesian IV estimation.

and interval estimators from ordinary least squares (OLS), 2SLS, and two variants of limited information maximum likelihood estimation (LIML) approaches, which are better suited estimators in case of weak or many instruments. Using artificial and real data, they claim that for “*both weak and strong instruments, our procedure produces credibility regions that are much smaller than competing classical procedures, particularly in the case of non-normal errors.*”

*Preview of results* In line with existing evidence on human capital outcomes, our 2SLS point estimates are relatively large, but statistically nonsignificant. In contrast, our semi-parametric Bayesian IV estimation provides precisely estimated zero effects. For instance, the estimated effect of family size on college attendance is 0.02 percentage points, with a corresponding 95 percent confidence interval ranging from minus 0.02 to 0.10. This confidence interval is about twenty-three times smaller than the equivalent confidence interval from the 2SLS estimation. For health outcomes, we also observe comparably better performance by the semi-parametric Bayesian IV estimate and do not find convincing evidence for an effect of family size. In summary, based on our empirical analysis, we conclude that there is *no* significant quantity-quality trade-off of children.

The remainder of this paper is organized as follows. Section 2 briefly outlines the conventional IV approach, and then presents both Bayesian IV approaches in more detail. Section 3 describes the data. Section 4 compares the empirical estimates based on the competing estimation methods. Finally, Section 5 concludes the paper.

## 2 Estimation Methods

### 2.1 Conventional Instrumental Variable Approach

Following the standard approach, we employ two complementary sources of exogenous variation in family size. The first source is twin births at the second live birth (Rosenzweig and Wolpin, 1980). The second source is same-sex sibships (Angrist and Evans, 1998), which exploits the widely observed preference among parents to have children of mixed sex. The first identification strategy (twin births) informs us about the effect of an unexpected third child, while the second (same-sex sibship) provides an estimate of the effect of an expected and desired third child. Both identification strategies are well-established natural experiments for studying the effect of family size on children’s human capital outcomes. They translate into the following two first-stage estimations,

$$fam\ size = \alpha^{T2} + \beta^{T2} \cdot twin2 + \Gamma^{T2} \cdot \mathbf{X} + u^{T2}, \quad (1)$$

$$fam\ size = \alpha^{SS} + \beta^{SS} \cdot same\ sex + \Gamma^{SS} \cdot \mathbf{X} + u^{SS}, \quad (2)$$

where *fam size* represents the completed family size (i.e., number of children). The respective estimation output (based on least squares) is summarized in Table 2.<sup>6</sup> The twin birth increases the family size for *all* families (with a twin birth at the second birth) who would have stopped fertility after the second live birth in the counterfactual situation of a single birth.<sup>7</sup> As can be seen in specification (1), the effect of a twin birth on family size is quite substantial, and amounts to about plus 0.78 children. On the other hand, the same-sex sibship only increases the family size for *some* families, namely those families (with a same-sex sibship) who decide, based on a preference for mixed siblings, to have a third child and who would have stopped fertility after the second live birth in the counterfactual situation of a mixed-sex sibship. Thus, this first stage relationship is comparably smaller, and amounts to only plus 0.06 children (see column (2)). Notably, compared to other papers, both first-stage relationships are quite large. This reflects the fact that Austrians have, on average, few children. For this reason, it is not worthwhile to use twin births at the third birth to examine the effect of a greater family size.<sup>8</sup>

An alternative specification of the same-sex sibship first-stage estimation allows for a different impact of female (*girl1, 2*) and male (*boy1, 2*) same-sex sibships. This specification (column (3)) reveals that the preference for a mixed sibling-sex composition is stronger in the case of two firstborn girls than it is if the two children are boys. In the former case, family size increases by 0.08 children, while in the latter case, it only increases by 0.04 children. This suggests that—apart from the general preference for mixed siblings—there is some prevalence of a preference for boys (Dahl and Moretti, 2008) among Austrian parents during that period. In our analysis below, we will employ specification (3) (instead of (2)) to exploit this additional variation. The lower panel of Table 2 displays the F-statistics on the excluded IV(s). It turns out that all our first-stage relationships are strong and that we can abstract from weak-IV concerns (Staiger and Stock, 1997) in the following discussion.<sup>9</sup>

For the twin-IV, we need to assume that the occurrence of twins is an exogenous event that is uncorrelated with unobserved parental characteristics. For the same-sex sibship IV, we have to make an equivalent assumption on the (order of) sex of the children. How plausible are these identifying assumptions? There are two known determinants of twin births. A higher maternal age and an *in vitro* fertilization treatment both increase the likelihood of multiple births. First, we control for maternal age. Second, we use

---

<sup>6</sup>Details on data sources will be provided below.

<sup>7</sup>This implies that there are no *never-takers* with this IV and the *local average treatment effect* (LATE) is equal to the *average effect on the non-treated* (ATC). See Angrist *et al.* (2010) for a more detailed discussion on this topic.

<sup>8</sup>The share of Austrian women with completed parity of four or more has declined since the 1935 birth cohort, and amounts to about 5 percent for the 1965 cohort (Prskawetz *et al.*, 2008).

<sup>9</sup>Depending on the outcome to be considered, the estimation samples vary. However, in each case, the first-stage relationships are comparable; the detailed first-stage estimation output for each sample is available upon request.

the occurrence of twins at the second birth, and assume that *in vitro* fertilization is less common when a woman has already given birth to a child. Furthermore, second births are predominantly from a period before *in vitro* fertilization treatment was customary. The upper graph in Figure 1 shows that, while the incidence of twin births has increased sharply in recent years (especially at parity one), a constant rate was maintained until the end of the 1980s. The bottom graph shows that in about 90 percent of our observations, second births took place before 1990. The child’s sex is typically assumed to be an outcome of a complete random biological process.<sup>10</sup> Thus, it also seems plausible to assume that a child’s sex is not (or is, but only marginally) correlated with unobserved family characteristics.

At the second stage of IV estimations, we use the two sources of exogenous variation in completed family size to identify the effect on different outcomes of the firstborn child. In particular, we estimate second-stage equations, where the outcome variable, *child outcome*, is either an educational, labor market, or health outcome measured in young adulthood,

$$child\ outcome = \alpha^{IV} + \beta^{IV} \cdot \widehat{fam\ size}^{IV} + \Gamma^{IV} \cdot \mathbf{X} + u^{IV}, \quad (3)$$

and  $\widehat{fam\ size}^{IV}$  is the completed family size instrumented by one of the two alternative strategies (indexed by *IV*) explained above. In our baseline specification, we include a parsimonious set of covariates (denoted by  $\mathbf{X}$ ) that comprises binary indicators of the firstborn’s year of birth and mother’s age at the first and second birth. The latter two variables capture the birth spacing.

## 2.2 Bayesian Instrumental Variable Approaches

Unlike the conventional approach, the Bayesian paradigm assumes that the parameters are random variables in addition to random observations. Consequently, distributions are specified for all random variables contained in the model and, by definition, the Bayesian approach relies far more on distributional assumptions than do conventional approaches, which typically include implicit rigid distributional assumptions (e.g., for estimating confidence intervals). More specifically, Bayesian inference requires formulating the conditional distribution of the data, given model parameters (i.e., the likelihood function), and a prior

---

<sup>10</sup>The so-called *Trivers-Willard Hypothesis* states that the population sex ratio responds to parental conditions (Trivers and Willard, 1973). It is predicted that mothers in good condition are expected to have more sons, while mothers in poor condition should have more daughters. The precise mechanism behind how mothers (or their reproductive system) favor either female or male offspring, depending on their condition, is still debated. As prenatal mechanisms, researchers discuss the prevention of the implantation of embryos of certain sex, or the increased likelihood of certain fetal loss. Empirical evidence is available that women in poor health (or who live in less favorable conditions) are less likely to have male offspring (Almond and Edlund, 2007; Catalano *et al.*, 2005; Hansen *et al.*, 1999). However, these effects are not quantitatively important to our results.



distribution for the model parameters themselves. This constitutes both a weakness and a strength of Bayesian modeling. On the one hand, the prior distributions of the model parameters are the main way of introducing information into the model that is not based on the data, such as the findings of comparable studies. Thus, choosing these distributions requires care. Typically, researchers prefer weakly informative prior distributions that do not conflict with the information provided by the data. However, the choice of such priors that are not based on real previous observations is, to a certain degree, debatable. On the other hand, Bayesian modeling allows for more flexibility when formulating the distribution of the data than do conventional approaches.

For Bayesian IV models, the typical approach concerning the prior distribution (as presented by Rossi *et al.* (2005)) is based on conjugate prior distributions, which preserve the structure of the distribution and allow simple parameter updates.<sup>11</sup> With regard to the likelihood function, two approaches have been suggested for Bayesian IV modeling, which we explain in detail below.

### 2.2.1 Standard Bayesian IV Approach

The standard Bayesian linear regression approach (Zellner, 1971; Rossi *et al.*, 2005; Robert, 2001), equivalent to classical OLS, is based on the second-stage equation, explicitly assuming a normal distribution for the error terms,

$$child\ outcome \sim N(\alpha + \beta \cdot fam\ size + \Gamma \cdot \mathbf{X}, \sigma_u^2). \quad (4)$$

The standard Bayesian approach towards inferring IV models can be seen as an extension of the standard Bayesian linear regression model, which explicitly assumes a bivariate normal distribution with a covariance matrix,  $Cov(u^{T2}, u^{IV}) = \Sigma$  for the error terms:

$$\begin{pmatrix} fam\ size \\ child\ outcome \end{pmatrix} = \begin{pmatrix} \beta^{T2} \cdot twin2 + \Gamma^{T2} \cdot \mathbf{X} \\ \beta^{IV} \cdot fam\ size^{IV} + \Gamma^{IV} \cdot \mathbf{X} \end{pmatrix} + \begin{pmatrix} u^{T2} \\ u^{IV} \end{pmatrix}, \quad (5)$$

$$\begin{pmatrix} u^{T2} \\ u^{IV} \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha^{T2} \\ \alpha^{IV} \end{pmatrix}, \Sigma\right). \quad (6)$$

---

<sup>11</sup>Kleibergen and Zivot (2003) compare classical and Bayesian estimators with such “non-informative” priors for IV models, focusing in particular on the specification of Bayesian models (i.e., likelihood and priors) that have similar properties to the 2SLS and LIML estimators. They find similarities between the choice of a flat prior over the parameters and the 2SLS estimator. They discuss the properties and advantages and disadvantages of the non-Bayesian and the Bayesian estimator in detail. Choosing a Jeffrey’s prior, which provides invariance with respect to parameter transformations, leads to the Bayesian equivalent of the LIML estimator. They find that Bayesian approaches are less sensitive to the inclusion of superfluous instruments, arguing that it pays to be aware of “implicitly assumed priors” in classical procedures.

A similar expression arises for the *same sex* IV. Following Conley *et al.* (2008), we included the intercepts  $\alpha^{T2}$ ,  $\alpha^{SS}$ ,  $\alpha^{IV}$  in all equations as the prior mean of the error terms to illustrate the relation between the standard and the semi-parametric Bayesian IV approach described below.

The prior distributions of the coefficients are joint normal distributions for all coefficient parameters,

$$\begin{pmatrix} \alpha^{T2} \\ \beta^{T2} \\ \Gamma^{T2} \end{pmatrix} \sim N(\mu_0^{T2}, \Sigma_0^{T2}), \quad (7)$$

with a similar prior,  $N(\mu_0^{SS}, \Sigma_0^{SS})$ , for the coefficients  $(\alpha^{SS}, \beta^{SS}, \Gamma^{SS})$  of the *same sex* IV. Analogously, the normal prior distribution,  $N(\mu_0^{IV}, \Sigma_0^{IV})$ , is assumed for the coefficients  $(\alpha^{IV}, \beta^{IV}, \Gamma^{IV})$ . Following Rossi *et al.* (2005), a conjugate inverse Wishart distribution<sup>12</sup> is assumed for the covariance matrix,  $\Sigma$ ,

$$\Sigma \sim IW(s_0, S_0). \quad (8)$$

We refer to this model as the standard Bayesian IV approach.

### 2.2.2 Semi-Parametric Bayesian IV Approach

The assumption of normally distributed errors is not realistic for most applied data settings. This also applies to our outcome variables, most of which are either binary or highly skewed (including excess zeroes), and our endogenous variables, which have small counts. Therefore, we suggest a semi-parametric approach based on a flexible Bayesian semi-parametric error model that is able to deal with this departure from normality. In particular, we follow Conley *et al.* (2008), who suggest a semi-parametric approach based on an infinite mixture of bivariate normal distributions as the error model:

$$\begin{pmatrix} u^{T2} \\ u^{IV} \end{pmatrix}, \begin{pmatrix} u^{SS} \\ u^{IV} \end{pmatrix} \sim \sum_{k=1}^{\infty} \pi_k N(\mu_k, \Sigma_k). \quad (9)$$

As in (6), the errors are not centered around zero, which enables us to identify all parameters of the infinite mixture (i.e., the weights  $\pi_1, \pi_2, \dots$ , which are normalized as  $\sum_{k=1}^{\infty} \pi_k = 1$ ), as well as the component-specific means and covariance matrices,  $\mu_k$  and  $\Sigma_k$ , for each component.

---

<sup>12</sup>The inverse Wishart distribution is defined as having the density

$$p(\Sigma) = \frac{|S_0|^{\nu/2}}{2^{\nu p/2} \Gamma_p(\nu/2)} |\Sigma|^{-(\nu+p+1)/2} \exp(-1/2 \text{tr}(S_0 \Sigma^{-1})).$$

In the special case of this model, only a single component has a corresponding weight equal to one, while all other components have zero weights. This corresponds to the standard Bayesian IV model, described in equation (6), with the mean,  $\mu$ , of the single non-empty component being equal to the intercept  $(\alpha^{T2}, \alpha^{IV})$ . More generally, a mixture of several components with non-zero weights could be regarded as a model in which the “overall” mean,  $\mu = \sum_{k=1}^{\infty} \pi_k \mu_k$ , of the error term in equation (9) is different from 0 and corresponds to a random intercept.

Hence, instead of specifying a prior for the intercepts  $(\alpha^{T2}, \alpha^{IV})$ , as in the standard Bayesian IV approach, a joint prior distribution has to be imposed on the countably infinite sequence of mixture component parameters,  $\mu_k, \Sigma_k$ , and their corresponding weight,  $\pi_k$ , for all  $k = 1, 2, \dots$ . A popular approach for specifying such a prior distribution is based on the notation of a discrete random probability measure, with the Dirichlet process prior introduced by Ferguson (1973) being an important special case. Under the Dirichlet process prior, the error distribution is represented as a countably infinite mixture, as in (9), where  $\pi_k$  are random weights, such that  $\sum_{k=1}^{\infty} \pi_k = 1$ . However, the components parameters,  $\theta_k = (\mu_k, \Sigma_k) \sim G_0$ , are independent draws from an underlying parametric family of distributions,  $G_0$ . The best way to understand the notion of a Dirichlet process prior is via the stick-breaking representation of the random weights, as introduced by Sethuraman (1994). This representation defines the infinite sequence  $(\pi_1, \pi_2, \dots)$  of random weights recursively from an infinite sequence of pieces of sticks  $(v_1, v_2, \dots)$ . These sticks are independent realizations from a beta distribution (i.e.,  $v_k \sim \mathcal{B}(1, \omega)$ ) with tightness parameter  $\omega$ . The idea behind “stickbreaking” is that we start with a stick of length 1, which is the sum of the weights. Then, we break the stick at length  $v_1$ , obtaining the first piece of stick,  $v_1$ , which defines the first weight,  $\pi_1 = v_1$ . This leaves the rest of the stick with the tail probability of  $1 - \pi_1$  for all remaining components. The second piece of stick is broken off from this remaining piece at length  $v_2$ , defining the second weight as  $\pi_2 = v_2(1 - v_1)$ . Once again, we are left with a piece of the stick, in this case with the tail probability of  $(1 - v_1)(1 - v_2)$  for all remaining components. This leads to the third weight,  $\pi_3 = v_3(1 - v_1)(1 - v_2)$ . As we continue to break the stick, weight  $\pi_k$  is defined as a piece with length  $v_k$ , and is taken from the remaining piece from the previous step with a tail probability of  $\prod_{j=1}^{k-1} (1 - v_j) = 1 - \sum_{j=1}^{k-1} \pi_j$ :

$$\pi_k = v_k \prod_{j=1}^{k-1} (1 - v_j). \quad (10)$$

This can be interpreted as the relative length of the k-th broken piece,  $v_k$ , broken off from the piece of stick remaining after breaking off  $v_1$  to  $v_{k-1}$ .

The distribution  $G_0$ , from which each  $\theta_k = (\mu_k, \Sigma_k)$  is drawn, is the so-called base measure, and is chosen as a conditionally conjugate prior for a model with a single com-

ponent. For the IV model defined in (5), this leads to choosing the following prior as a base measure for each component parameter,  $\theta_k = (\mu_k, \Sigma_k)$ :

$$\Sigma_k \sim IW(\nu, V), \quad \mu_k | \Sigma_k \sim N(\bar{\mu}, a^{-1} \Sigma_k). \quad (11)$$

Here the choice of prior parameters  $\nu$  and  $V$  influences the prior amount of shrinkage toward the case of no endogeneity.

Although the semi-parametric Bayesian IV model assumes, a priori, a mixture of infinitely many components, for a finite number of realizations from this model, the errors are classified only into finitely many non-empty mixture components. Each of these components is characterized by a distinct value of  $(\mu_k, \Sigma_k)$ , for  $k = 1, \dots, I^*$ , where  $I^*$  is the number of non-empty, distinct mixture components.

The prior number of non-empty, distinct mixture components,  $I^*$ , is random by construction and strongly depends on the choice of the tightness parameter,  $\omega$ . In this respect, it is useful to consider representation (10), which shows that the  $k$ th piece of stick  $v_k$  defines the fraction of the tail probability,  $\prod_{j=1}^{k-1} (1 - v_j)$ , assigned to the  $k$ th component. Since the expected stick length is equal to  $E(v_k) = 1/(1 + \omega)$ , choosing a small value for  $\omega$  leads, a priori, to fewer components,  $I^*$ , and vice versa. To reduce this sensitivity to the choice of  $\omega$ , a hierarchical prior is added for  $p(\omega)$ .

Finally, we assume normal prior distributions for the remaining coefficients, as in equation (7). In other words:

$$\begin{pmatrix} \beta^{T2} \\ \Gamma^{T2} \end{pmatrix} \sim N(\mu_0^{T2}, \Sigma_0^{T2}). \quad (12)$$

Once again, the priors for  $(\beta^{SS}, \Gamma^{SS})$  and  $(\beta^{IV}, \Gamma^{IV})$  are defined analogously. We refer to this model as the semi-parametric Bayesian IV approach.

### 2.2.3 Bayesian Inference

For both Bayesian IV approaches, we use Markov Chain Monte Carlo (MCMC) methods for Bayesian posterior inference. Since the standard Bayesian IV model is estimated under a bivariate normal distribution for the error vector  $(u^{T2}, u^{IV})$ , a three-block Gibbs sampler can be applied for inference (see Rossi *et al.* (2005) and Hoogerheide *et al.* (2007) for more information). Based on the priors defined in (7) and (8), a sampler is applied that updates the coefficient parameters of the first-stage equations, the second-stage equations, and the covariance matrix of the errors.

Similarly, the Bayesian posterior inference for the semi-parametric Bayesian IV model is carried out using an MCMC estimation (see Conley *et al.* (2008) for details). In contrast to the previous method, the coefficients  $(\beta^{T2}, \Gamma^{T2})$ ,  $(\beta^{SS}, \Gamma^{SS})$ , and  $(\beta^{IV}, \Gamma^{IV})$  are

sampled jointly in different blocks, conditional on all remaining parameters. The mixture components are updated in a further block, using the Polya urn representation to draw the corresponding component-specific parameter,  $\theta_i$ , for each observation  $i$ , conditional on the parameters of all remaining observations (e.g., see (Escobar and West, 1998)). The Polya urn scheme is an alternative representation to the Dirichlet process prior. For each observation  $i$ , this scheme provides the probability that  $\theta_i$  is equal to one of the  $I^*$  distinct parameters of the remaining observations, as well as the probability that  $\theta_i$  belongs to a new, empty component, in which case  $\theta_i$  is sampled from the base measure,  $G_0$ . This algorithm automatically increases or decreases the number of non-empty, distinct components,  $I^*$ , during MCMC sampling. Hence, a posteriori, the number of non-empty, distinct mixture components,  $I^*$ , in (9) is adjusted to the number of mixture components required to best model the non-normality of the data.

The MCMC estimation for both Bayesian IV approaches is implemented in the R package, `bayesm` (see Rossi (2012)), designed for the statistical software environment, R (R Core Team, 2013). The large calculation times of the Dirichlet Process sampler prohibits an extensive sensitivity analysis of the prior settings. Therefore, we applied the well-defined default priors of `bayesm`. However, we encountered problems with the Dirichlet process sampler implemented in `bayesm`, as the algorithm was not able to manage the large data sets we analyze in this study. This obstacle prohibited a straightforward application of the given software. Thus, we developed an alternative approach by splitting the various data sets into a dozen smaller data sets,  $D_1, \dots, D_K$ . Then, we ran the MCMC estimation on each data set,  $D_k$ , and performed a post-inference resampling step to sequentially merge the  $K$  posterior densities.

### 3 Data

The empirical analysis is based on several administrative data sources from Austria. In the *Austrian Birth Register*, we select all firstborn single births from the birth cohorts 1979 through 1983. Of these, we focus on mothers between 18 and 40 years of age at the time of their first birth, and who had at least one subsequent birth. This implies that the mothers in our sample belong to the birth cohorts from 1939 through 1965. Based on the characteristics of the second birth, we define our IVs, *twin2* and *same sex*. Here, we exclude the rare cases of higher-order multiples. To generate our treatment variable, *fam size*, we add all the children born until 2007. Given that the youngest mother in our sample was 42 years old in 2007, this variable can be interpreted as completed fertility. Figure 2 shows the distribution of the firstborns' year of birth, completed fertility size, mothers' age at first birth, and the years between first and second birth (birth spacing). Information on the covariates is also captured from the birth register. Finally, to explore potentially heterogeneous effects across households with a varying socioeconomic

status, we use available information on mothers' educational attainment. Using a personal identifier, the data from the birth register are matched to other data sources comprising information on different firstborns' outcomes.<sup>13</sup>

*Educational outcome:* We do not have complete information on firstborns' educational attainment. However, a link to the database of the *Austrian Federal Ministry of Labour, Social Affairs and Consumer Protection*, which includes information on family allowances, school and study grants, or other types of public social benefits (until 2005), allows us to infer whether firstborns were ever enrolled in university. Given the Austrian educational system, this is an informative proxy for educational attainment, and can be interpreted as the completion of (at least) the so-called *high track*.<sup>14</sup> The binary outcome variable, "college attendance," has a mean of about 0.3.

*Labor market outcomes:* To evaluate firstborns' labor market performance, we use the *Austrian Social Security Database (ASSD)*. This is an administrative record used to verify pension claims for the universe of Austrian workers. Here, we observe firstborns' employment history (including basic employer information), unemployment, and various other qualifications on a daily basis. Information on earnings is provided per year and per employer. The limitations of the data are top-coded wages and the lack of information on working hours (Zweimüller *et al.*, 2009). For our analysis, we evaluate employment, occupation, and wages. All outcome variables are measured in the last quarter of the year 2011, which is when individuals are between 28 and 32 years old. At this time, 80 percent of the firstborns are employed (71 percent as white-collar workers and 29 percent as blue-collar worker). Their average daily wage is about € 84.

*Health outcomes:* Health outcomes are only available for the sub-population of all private sector employees and their dependents in the province of Upper Austria.<sup>15</sup> We have access to the database of the *Upper Austrian Sickness Fund*<sup>16</sup>, which includes detailed information on healthcare service utilization in the outpatient sector (i.e., medical attendance and drug use) and some inpatient sector information, such as the number of days of

---

<sup>13</sup>This means that, in contrast to other papers (Vere, 2011), we have exact information on multiple births. Therefore, we do not have to infer this information from the year of birth. Thus, the matched information from the *Austrian Birth Register* eliminates problems of measurement error.

<sup>14</sup>The Austrian educational system is characterized by two distinguishing features: *early tracking* and a widespread dual education system. Students are already allocated in Grade five (i.e., at the age of 10) to one of two different educational tracks. The lower secondary schools (*low track*) comprise Grades 5 to 8, provide basic general education and prepare students for vocational education either in an intermediate vocational/technical school or within the dual education system. The higher secondary schools (*high track*) comprise a first-stage (Grades 5 to 8) and second stage (Grades 9 to 12), provide advanced general education, and conclude with a university entrance exam. The majority of the students (about 72 percent) complete the *low track*, and only about 30 percent complete the *high track*.

<sup>15</sup>Upper Austria is one of nine provinces in Austria and comprises about one sixth of the Austrian population and work force.

<sup>16</sup>In Austria, sickness funds cannot be freely chosen by the insured. Membership is mandatory, and the assignment of employees to a particular fund depends on the location of the employer and on the type of occupation. The more than one million fund members represent approximately 75 percent of the Upper Austrian population.

hospitalization and the Sickness Fund’s financial contribution for a patient’s hospital stay. For instance, we are able to observe each single doctor visit and each drug prescription, together with the exact date of service utilization. To evaluate health, we construct the following annual outcome variables: (i) total health expenditure<sup>17</sup>; (ii) expenditure on outpatient medical attendance (at general practitioners (GPs) and resident medical specialists); (iii) expenditure on medical drugs; and (iv) days of hospitalization. Obviously, the degree to which these proxies reflect individual health varies among the variables. Whereas the number of days of hospitalization and the consumption of medical drugs can be expected to be highly correlated with a person’s health status, expenditure on outpatient medical attendance may also capture aspects of preventative care, such as costs of screening exams. Therefore, higher outpatient expenditures do not necessarily reflect poorer health. All variables are measured in the year 2009, when firstborns are between 26 and 30 years old. The average firstborn spent about €584 on total health expenditure (€219 on outpatient medical attendance and €87 on medical drug consumption) and spent 1.1 days in hospital.

## 4 Estimation Results

### 4.1 Human Capital Outcomes

Table 3 summarizes our estimation results with regard to the effect of family size on firstborns’ educational and labor market outcomes. Column (1) provides the sample mean of the respective outcome variable. The OLS estimates in column (2) suggest a statistically significant negative correlation between family size and all human capital outcomes. An increase of one sibling is associated with a reduced likelihood of college attendance (minus 1.2 percentage points) and a lower employment probability (minus 1.7 percentage points). Conditional on being employed, an additional sibling reduces the likelihood of white-collar employment (as compared to blue-collar employment) by 3.3 percentage points and also has a negative effect on the wage. The latter effect amounts to a reduction in the daily wage of about €1.7, a decrease of 2 percent. The estimates obtained by an equivalent Bayesian method (see column (3)) are quantitatively comparable.

Columns (4) and (7) summarize the 2SLS estimation results for the two different IVs. With one exception, statistical significance vanishes. While some 2SLS estimates are within a similar range to comparative OLS estimates, the standard errors are consistently higher. On average, the 95 percent confidence intervals for the twin IV estimates are 10 times larger than those of the OLS estimation. The corresponding intervals from the same-sex sibship IV are more than 23 times larger than their OLS counterparts. In the

---

<sup>17</sup>Total health expenditures cover outpatient expenditures plus a fixed fee paid by the Sickness Fund per day spent in hospital.

case of the twin IV, the point estimates for college attendance and employment are still negative, while we observe a positive sign for white-collar employment and wage. The only statistically significant impact of family size is the one on white-collar employment, based on the same-sex sibship IV. Here, one additional sibling increases the likelihood of working as a white-collar worker by 9.3 percentage points.

Columns (5) and (8) in Table 3 summarize the estimation results based on the standard Bayesian IV approach. Below these estimates, we provide highest posterior density intervals (HPDI). These intervals cover 95 percent of the posterior distribution and are located in such a way that they provide the shortest interval. For both instruments, we find comparable results to those of the simple Bayesian estimations, both with respect to point estimates and confidence intervals. Family size is estimated to have a statistically significant negative effect on college attendance, white-collar employment, and wages. However, the effect on employment is positive and also significant.

Both the standard Bayesian and standard Bayesian IV regressions rely on a normality assumption. However, our data clearly violates the normal distribution assumption of these models, as we are faced with mainly binary outcomes and a discrete exogenous variable. The strength of the semi-parametric approach lies in approximating non-normal distributions, even complex ones, using a mixture of normal distributions. The large number of components required for this mixture can be viewed as an indication of the approximated distribution being very dissimilar to the normal distribution. In order to check the degree of “non-normality,” we express the distribution of the number of components using the mixture algorithm required in Table 6. We observe that the smallest number of components is required for the dichotomous labor outcomes, “college attendance,” “employment,” and “white-collar,” which are discrete and non-normal. However, they can be represented by the mixture with relatively few components. The median 4 – 5 components suffice to model the binary outcomes’ and correlated first-stage equation’s joint error distribution. Modeling the corresponding error distribution for the skewed and censored wages requires more components for the median 7 – 8 components.

Columns (6) and (9) in Table 3 summarize the estimation results from the semi-parametric Bayesian IV approach, which allows for departure from normality. In the case of the twin IV, statistically significant effects are no longer visible. The size of the estimates is almost negligible, and with the exception of daily wages, essentially zero. The HPDIs for the outcomes of college attendance, employment, and white-collar employment lie in a small range between minus 0.3 and plus 0.1 percentage points. The nonsignificant estimate for daily wage is €2.72, which is equivalent to an increase of 3.2 percent. In the case of the same-sex sibship IV, two estimates are statistically significant and negative (college attendance and white-collar employment), although their economic significance of minus 0.35 and plus 0.05 percentage points is negligible. The still nonsignificant coefficient for daily wages indicates a reduction of 4 percent.



Figure 3 provides a graphical summary of the IV estimates for two outcomes and compares them to estimates from Angrist *et al.* (2010), who use data from the 1983 and 1995 Israeli censuses (linked with information from the population registry). The panels depict the statistically nonsignificant 2SLS estimates, with large confidence intervals, covering economically important outcome ranges. We argue that these estimates are not precise enough to reject the trade-off between the quantity and quality of children. The estimates provided by the standard Bayesian IV approach are substantially more precise and statistically significant throughout. However, they do not provide a consistent picture. For example, they suggest that both an unexpected and expected increase in family size have a negative effect on firstborns' college attendance, but a positive effect on their employment probability. While this result is theoretically possible—for instance, an exogenous increase in family size may prevent a firstborn from picking a college major with a poor employment outlook—we have to keep the imposed normality assumption in mind. Finally, the estimates of the semi-parametric Bayesian IV approach consistently show either statistically and/or economically nonsignificant effects. We argue that these estimation results are precise enough to credibly reject an (economically relevant) interaction between the quantity and quality of children with respect to human capital outcomes.

## 4.2 Health Outcomes

For our health outcomes, we estimate an equivalent set of models, as summarized in Table 4. Across all models and outcomes, we find consistently nonsignificant effects of family size.<sup>18</sup> In comparison to labor market outcomes, the size of the different confidence intervals provides a more differentiated picture. For 2SLS, the estimates are again too imprecise to be informative. The estimations based on the twin IV give confidence intervals are, on average, 6 times larger than those of the OLS estimation. This factor goes up to 25 in the case of same-sex sibship IV. As an extreme, with a mean of 1.11 days of hospitalization, the 95 percent confidence interval for the same-sex sibship lies in the range minus 3 to plus 2.5 days (see column (7)).

The standard Bayesian IV approach provides quite precise estimates. However, the objection against this method (the violation of the normality assumption) is even more relevant for health outcomes. Next, we view a more complex scenario than that of the predominantly binary human capital outcomes. The cardinal health outcome variables exhibit a skewed distribution, since a large fraction of those insured do not utilize any medical services and/or drugs in a given year. Consequently, we cannot rely on either the

---

<sup>18</sup>The only two exceptions are the estimates obtained from the standard Bayesian regression (column (3)) for drug expenditure and days spent in hospital. The effect of an increase in family size on days of hospitalization is negligible, while the negative association with the expenditure on drug consumption is equivalent to a decrease of 5.1 percent. However, the HDPI regions of the standard Bayesian regression are based on the assumption of homoscedastic errors, whereas the confidence regions reported by the OLS technique allow for heteroscedasticity of unknown form.

point or the variance estimates from the standard Bayesian IV approach, and we consider the semi-parametric method to provide the most reliable inference results.

Table 6 compares the number of components required by the semi-parametric Bayesian IV approach. As expected, the number of components required for health outcomes is substantially higher (with a median between 10 and 21) than the human capital outcomes (with a median between 4 and 8). The difference shows the difficulty of approximating the true error distribution. This results in comparable larger HPDIs for our health outcomes (compared to those of the human capital outcomes). However, for the twin IV, the performance is still very good. The point estimates are reasonably close to zero, and the HPDIs are small; at least substantially smaller than those of the corresponding 2SLS estimations. In the case of the same-sex sibship IV, we increase the endogenous variable's variation in the first stage equation by adding a second instrument (*girl1, 2* and *boy1, 2*) and boosting the variation in the posterior distribution. Thus, even though the complexity of the mixture, represented by its number of components in Table 6, does not increase, the variation of the posterior distribution does. Consequently, the precision declines in the case of the same-sex sibship IV estimates. While the estimates are still more precise than those of the 2SLS estimation (with a relative order of magnitude of about one half), we still cannot reject the quantity-quality trade-off or the existence of the hygiene hypothesis.

To put our zero results for health outcomes into perspective, it is worth noting that the available causal evidence is also insignificant, economically or statistically (see Table 1). As was mentioned earlier, Hatton and Martin (2010) find a negative effect on children's health, although the impact on body weight remains nonsignificant. Apart from the fact that the external validity of the study is at least debatable (the sample consists of predominantly poor families in Britain taken in the interwar period from 1937–1939), the quantitative effect is only 1 percent, which does not seem to be very high, given the degree of negative economic selection of families. Similarly, Lundborg *et al.* (2013) find a significant positive effect of family size on children's health, although the quantitative significance (0.4 percent) is even lower than that of Hatton and Martin (2010). Lastly, an increase in the number of children by one improves the global health of males at the age of 18 by 4 percent. This effect is obviously quantitatively important, although the estimated coefficient is only significant at the 10 percent level.

### 4.3 Heterogenous Treatment Effects

How family size affects firstborns' outcomes may depend on specific circumstances. In particular, the socioeconomic background of families may matter. The degree to which financial resources are binding and to which parents are prepared to provide non-pecuniary support to an additional child is expected to vary across the socioeconomic spectrum. There is some evidence in the literature (based on 2SLS estimates) that a quantity-quality

trade-off is significant among families from a lower socioeconomic background, but is non-existent among families with a higher socioeconomic status (Åslund and Grönqvist, 2010). If this also applies to our data (and carries over to our preferred estimation model), the examination of average treatment effects (as presented above) may hide important aspects. To identify potential heterogeneous treatment effects, we stratify our sample by the mothers' educational attainment measured at the time of their second birth. We distinguish between mothers with three different levels of educational attainment. Mothers in the first group have compulsory schooling only (about 27 percent), those in the second group graduated from the low track (about 58 percent), and those in the third group graduated from the high track (about 15 percent).

We present the estimation results from these stratified samples for human capital outcomes only, as the resulting number of observations is sufficient. The semi-parametric Bayesian IV estimates in columns (6) and (9) in Table 5 do not indicate remarkable differences between the socioeconomic groups. Similar to the aggregated results, the majority of effects are still nonsignificant, and the zeroes are precisely estimated. However, we do observe single statistically significantly negative effects for the group with compulsory schooling only, as well as the graduates from the low track. For children from mothers who completed compulsory schooling only, we find a negative effect on the employment probability of minus 0.5 percentage points in the same-sex sibship specification (column (9)), which is equivalent to a reduction of 0.6 percent. For children from mothers who graduated from the low track, we observe a negative effect on the likelihood of college attendance of minus 0.5 percentage points (2 percent) for the twin IV (column (6)), and of minus 0.2 percentage points in the same-sex sibship IV framework (column (9)). Hence, the stratified analysis confirms our result that an exogenous increase in family size has no significant effects on firstborns' human capital outcomes. If we find single significant effects, they are of minor economic importance. However, the detrimental effects of larger families on educational and labor market success of offspring, in accordance with the model for the quantity and quality of children, occur in families with a lower socioeconomic status.

## 5 Conclusions

In this study, we used a semi-parametric Bayesian IV approach to test the theoretically proposed interaction between the quantity and the quality of children. We used administrative data sources from Austria and exploited exogenous variation in family size due to twin births and preferences for a mixed sibling-sex composition. In contrast to existing evidence (based on 2SLS), our approach provides precisely estimated effects. By construction, the Bayesian semi-parametric approach is able to cope well with data that violate the typical normal distribution assumptions, such as binary, skewed, and censored data,

all of which occur in our data set. This allows us to credibly conclude that there is no quantitatively important quantity-quality trade-off of children. The estimated effects on firstborns' human capital and health outcomes are either statistically and/or economically nonsignificant.

How can the absence of this causal link be explained? There are several complementary explanations. First, the negative effect of reduced resources might be offset by the positive effects from interaction between siblings. This is what the hygiene hypothesis suggests for health outcomes. In terms of human capital development, firstborns may benefit from teaching their younger sibling. Second, parents may manage to keep their investment in their children constant (as compared to the counterfactual situation) and primarily reduce their consumption of private goods and leisure. Related literature has shown that mothers reduce their labor supply as a response to fertility shocks. Consequently, the total amount of at-home care provided by the mother may even increase for the firstborns. If maternal childcare is superior to the counterfactual care setting, firstborns might benefit from this mediating channel. Finally, an increase in family size might only reduce irrelevant child investment. In other words, the parents may overinvest in their firstborns in the counterfactual situation.

Naturally, our results have to be interpreted in the light of the specific institutional setting. We consider Austria to be a representative example of a Central European welfare state that provides—apart from free healthcare and education—some financial support for families. Moreover, the average family size in Austria is low, and our results relate basically to an increase in family size from two to three children. It remains unclear whether the quantity/quality trade-off would exist in developing countries with (i) a less pronounced welfare state and (ii) larger families, on average. In any case, it would be worthwhile to further expand the outcome dimension, and to study, for instance, the effect of family size on children's cognitive and character skills. These outcomes have recently attracted significant attention from economists, as there is growing evidence that these skills are strong predictors of important long-term outcomes (Heckman and Kautz, 2012, 2013).

## References

- ALMOND, D. and EDLUND, L. (2007). Trivers-Willard at Birth and One Year: Evidence from US Natality Data 1983-2001. *Proceedings of the Royal Society B: Biological Sciences*, **274** (3), 2491–2496.
- 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), 773–823.
- ANGRIST, J. D. and EVANS, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, **88** (3), 450–477.
- ÅSLUND, O. and GRÖNQVIST, H. (2010). Family Size and Child Outcomes: Is There Really No Trade-off? *Labour Economics*, **17** (1), 130–139.
- BAGGER, J., BIRCHENALL, J. A., MANSOUR, H. and URZUÁ, S. (2013). Education, Birth Order, and Family Size, unpublished manuscript.
- BECKER, G. S. and LEWIS, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, **81** (2), 279–288.
- and TOMES, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, **84** (4), 143–162.
- 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. *Quarterly Journal of Economics*, **120** (2), 669–700.
- , — and — (2010). Small Family, Smart Family? *Journal of Human Resources*, **45** (1), 33–58.
- BRINCH, C. N., MOGSTAD, M. and WISWALL, M. (2012). Beyond Late with A Discrete Instrument. Heterogeneity in the Quantity-Quality Interaction of Children. *Discussion Papers No. 703, September 2012, Statistics Norway, Research Department*.
- CÁCERES-DELPIANO, J. (2006). The Impacts of Family Size on Investment in Child Quality. *Journal of Human Resources*, **41** (4), 738–754.
- CATALANO, R., BRUCKNER, T., ANDERSON, E. and GOULD, J. B. (2005). Fetal Death Sex Ratios: A Test of the Economic Stress Hypothesis. *International Journal of Epidemiology*, **34**, 944–948.
- CHAMBERLAIN, G. and IMBENS, G. (2003). Nonparametric Applications of Bayesian Inference. *Journal of Business & Economic Statistics*, **21** (1), 12–18.
- CONLEY, T., HANSEN, C., MCCULLOCH, R. and ROSSI, P. (2008). A Semi-Parametric Bayesian Approach to the Instrumental Variable Problem. *Journal of Econometrics*, **144** (1), 276–305.
- DAHL, G. B. and MORETTI, E. (2008). The Demand for Sons. *Review of Economic Studies*, **75** (4), 1085–1120.

- ESCOBAR, M. D. and WEST, M. (1998). Computing Nonparametric Hierarchical Models. In D. Dey, P. Müller and D. Sinha (eds.), *Practical Nonparametric and Semiparametric Bayesian Statistics*, no. 133 in Lecture Notes in Statistics, Berlin: Springer-Verlag, pp. 1–22.
- FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, **1**, 209–230.
- FORASTIERE, F., AGABITI, N., CORBO, G. M., DELL’ORCO, V., PORTA, D., PISTELLI, R., LEVENSTEIN, S. and PERUCCI, C. A. (1997). Socioeconomic Status, Number of Siblings and Respiratory Infections in Early Life as Determinants of Atopy in Children. *Epidemiology*, **8** (5), 566–570.
- HANSEN, D., MØLLER, H. and OLSEN, J. (1999). Severe Periconceptional Life Events and the Sex Ratio in Offspring: Follow Up Study Based on Five National Registers. *British Medical Journal*, **319**, 548–549.
- HATTON, T. J. and MARTIN, R. M. (2010). The Effects on Statute of Poverty, Family Size, and Birth Order: British Children in the 1930s. *Oxford Economic Papers*, **62** (1), 157–184.
- HECKMAN, J. J. and KAUTZ, T. (2012). Hard Evidence on Soft Skills. *Labour Economics*, **19** (4), 451–464.
- and — (2013). *Fostering and Measuring Skills: Interventions That Improve Character and Cognition*. Discussion Paper 7750, Institute for the Study of Labor, Bonn, Germany.
- HOOGERHEIDE, L., KLEIBERGEN, F. and VAN DIJK, H. (2007). Natural Conjugate Priors for the Instrumental Variables Regression Model Applied to the Angrist-Krueger Data. *Journal of Econometrics*, **138**, 63–103.
- JARVIS, D., CHINN, S., LUCZYNSKA, C. and BURNEY, P. (1997). The Association of Family Size with Atopy and Atopic Disease. *Clinical and Experimental Allergy*, **27** (3), 240–245.
- KARMAUS, W. and BOTEZAN, C. (2002). Does a Higher Number of Siblings Protect Against the Development of Allergy and Asthma? A Review. *Journal of Epidemiology and Community Health*, **56** (3), 209–217.
- KLEIBERGEN, F. and ZIVOT, E. (2003). Bayesian and Classical Approaches to Instrumental Variable Regression. *Journal of Econometrics*, **114**, 29–72.
- LOPES, H. F. and POLSON, N. G. (2014). Bayesian Instrumental Variables: Likelihoods and Priors. *Econometric Reviews*, **33** (1–4), 100–121.
- LUNDBORG, P., RALSMARK, H. and ROTH, D. (2013). *When a Little Dirt Doesn’t Hurt – The Effect of Family Size on Child Health Outcomes*. Working paper, Linnaeus University Centre for Labour Market and Discrimination Studies.
- MOGSTAD, M. and WISWALL, M. (2012). Testing the Quantity-Quality Model of Fertility: Linearity, Marginal Effects and Total Effects. *mimeo*.

- POLLET, T. V. and HOBEN, A. D. (2011). An Evolutionary Perspective on Siblings: Rivals and Resources. In S. C.A. and T. K. Shackelford (eds.), *Handbook of Evolutionary Family Psychology*, Oxford University Press, pp. 128–148.
- PRSKAWETZ, A., SOBOTKA, T., BUBER, I., ENGELHARDT, H. and GISSER, R. (2008). Austria: Persistent Low Fertility Since the Mid-1980s. *Demographic Research*, **19** (12), 293–360.
- QUIAN, N. (2013). Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China. *mimeo*.
- R CORE TEAM (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- ROBERT, C. P. (2001). *The Bayesian Choice*. Springer.
- ROSENZWEIG, M. R. and WOLPIN, K. I. (1980). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*, **48** (1), 227–240.
- ROSSI, P. (2012). bayesm: Bayesian Inference for Marketing/Micro-econometrics. R package version 2.2-5.
- ROSSI, P. E., ALLENBY, G. M. and MCCULLOCH, R. (2005). *Bayesian Statistics and Marketing*. Chichester: John Wiley and Sons.
- SETHURAMAN, J. (1994). A Constructive Definition of Dirichlet Priors. *Statistica Sinica*, **4**, 639–650.
- STAIGER, D. and STOCK, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65** (3), 557–586.
- STRACHAN, D. P. (1989). Hay Fever, Hygiene, and Household Size. *British Medical Journal*, **299** (6710), 1259–1260.
- (2000). Family Size, Infection and Atopy: The First Decade of the ‘Hygiene Hypothesis’. *Thorax*, **55** (Suppl 1), S2–S10.
- , HARKINS, L. S., JOHNSTON, I. A. and ANDERSON, H. (1997). Childhood Antecedents of Allergic Sensitization in Young British Adults. *Journal of Allergy and Clinical Immunology*, **99** (1), 6–12.
- SVANES, C., JARVIS, D., CHINN, S. and BURNEY, P. (1999). Childhood Environment and Adult Atopy: Results from the European Community Respiratory Health Survey. *Journal of Allergy and Clinical Immunology*, **103** (3), 415–420.
- TRIVERS, R. L. and WILLARD, D. E. (1973). Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring. *Science*, **179** (4068), 90–92.
- VERE, J. P. (2011). Fertility and Parents’ Labour Supply: New Evidence from US Census Data. *Oxford Economic Papers*, **63** (2), 211–231.
- VON MUTIUS, E., MARTINEZ, F. D., FRITZSCH, C., NICOLAI, T., REITMEIR, P. and THIEMANN, H. (1994). Skin Test Reactivity and Number of Siblings. *British Medical Journal*, **308** (6930), 682–695.

- WICKENS, K., CRANE, J., PEARCE, N. and BEASLEY, R. (1999). The Magnitude of the Effect of Smaller Family Sizes on the Increase in the Prevalence of Asthma and Hay Fever in the United Kingdom and New Zealand. *Journal of Allergy and Clinical Immunology*, **104** (3), 554–558.
- WILLIS, R. J. (1973). A New Approach to the Economic Theory of Fertility Behavior. *Journal of Political Economy*, **81** (2), 14–64.
- ZELLNER, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons.
- , ANDO, T., BASTURK, N., HOOGERHEIDE, L. and VAN DIJK, H. (2012). *Bayesian Analysis of Instrumental Variable Models: Acceptance-Rejection within Direct Monte Carlo*. Tinbergen Institute Discussion Papers 12-098/III, Tinbergen Institute.
- ZWEIMÜLLER, J., WINTER-EBMER, R., LALIVE, R., KUHN, A., WUELLRICH, J., RUF, O. and BÜCHI, S. (2009). *The Austrian Social Security Database (ASSD)*. Working Paper 0901, The Austrian Center for Labor Economics and the Analysis of the Welfare State, University of Linz.



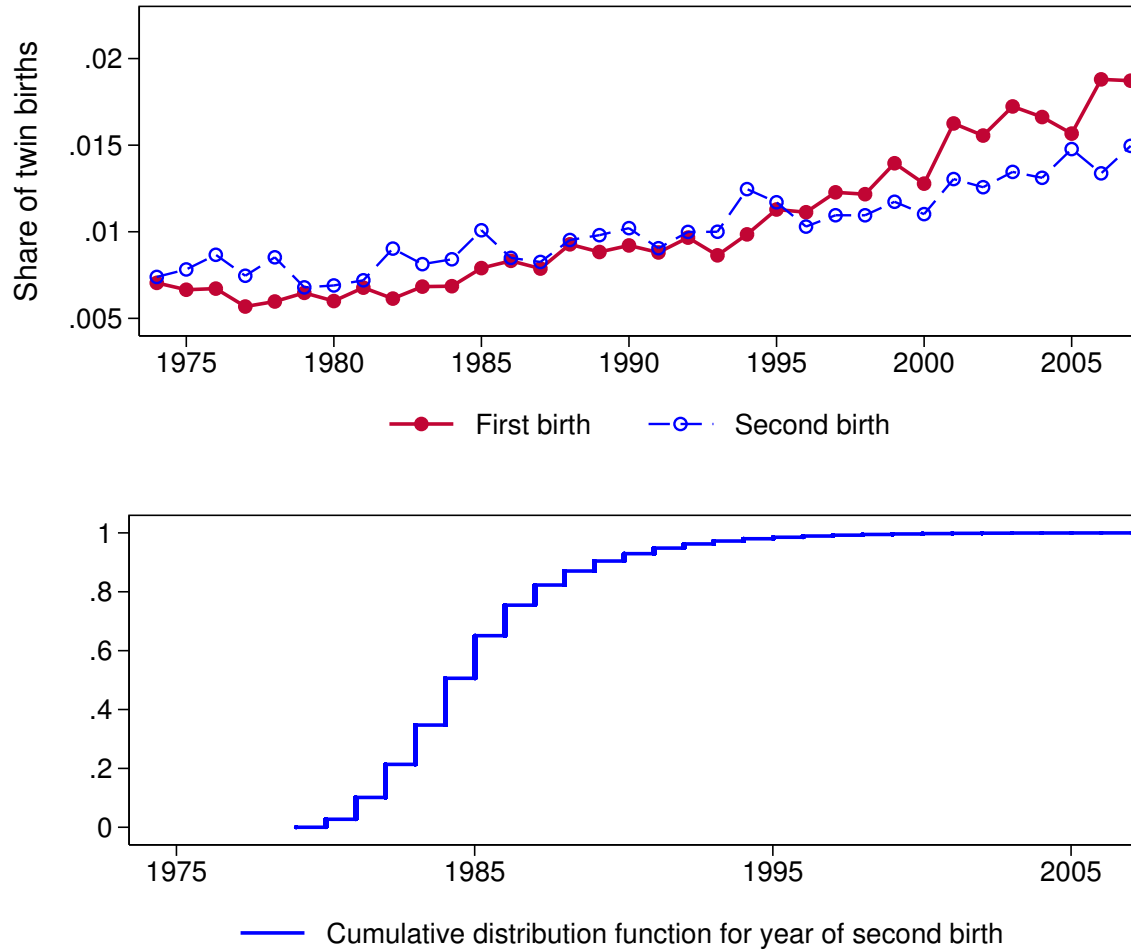
# Tables and Figures

**Table 1: 2SLS-Estimates of Family Size on Human Capital and Health Outcomes in the Literature**

Outcome	Mean	2SLS estimate	Standard error	95 percent confidence interval	
<b>Angrist <i>et al.</i> (2010)<sup>a</sup></b>					
Highest grade completed	12.6	0.174	0.166	-0.151	0.499
Years of schooling $\geq 12$	0.824	0.03	0.028	-0.025	0.085
Some college (age $\geq 24$ )	0.291	0.017	0.052	-0.085	0.119
College graduate (age $\geq 24$ )	0.202	-0.021	0.045	-0.109	0.067
Employment	0.827	-0.005	0.038	-0.079	0.069
Hours worked last week	32.6	-0.97	2.58	-6.03	4.09
Monthly earnings (1995 shekels)	2,997	-7.7	394.1	-780.1	764.7
Log of earning for full-time workers	8.24	0.082	0.116	-0.145	0.309
<b>Black <i>et al.</i> (2005)<sup>b</sup></b>					
Years of schooling	12.2	0.038	0.047	-0.054	0.130
Full-time employment (men)	0.71	-0.005	0.013	-0.030	0.020
Full-time employment (women)	0.46	0.021	0.015	-0.008	0.050
Log of earnings (men)	12.55	-0.003	0.02	-0.04	0.04
Log of earnings for full-time workers (men)	12.72	-0.005	0.014	-0.032	0.022
Log of earnings (women)	12.05	0.03	0.026	-0.021	0.081
Log of earnings for full-time workers (women)	12.42	0.015	0.016	-0.016	0.046
<b>Black <i>et al.</i> (2010)<sup>b</sup></b>					
IQ (stanine scores)	5.2	-0.149	0.052	-0.251	-0.047
Years of schooling $\geq 12$	12.3	-0.029	0.015	-0.058	0.000
<b>Åslund and Grönqvist (2010)<sup>a</sup></b>					
Years of schooling	12.9	0.022	0.048	-0.072	0.116
Grade point average (compulsory school)	51.7	0.844	0.665	-0.459	2.147
Grade point average (secondary school)	51.37	0.969	0.746	-0.493	2.431
Graduated from secondary school	0.92	0.002	0.007	-0.012	0.016
Enrolled in university	0.47	0.005	0.011	-0.017	0.027
Non-employment	0.21	-0.007	0.01	-0.03	0.01
Log of earnings	7.2	0.044	0.03	-0.01	0.10
<b>Cáceres-Delpiano (2006)<sup>a</sup></b>					
Grade behind cohort <sup>c</sup>	0.05	0.002	0.003	-0.004	0.008
Attends private school	0.15	-0.012	0.005	-0.022	-0.002
<b>Hatton and Martin (2010)<sup>d</sup></b>					
Height at eight [in cm]	121.00	-1.20			
<b>Lundborg <i>et al.</i> (2013)<sup>a</sup></b>					
Height of 18 year old males	179.691	1.102	0.360	0.396	1.808
Global health of 18 year old males <sup>e</sup>	-2.35	0.094	0.055	-0.014	0.202

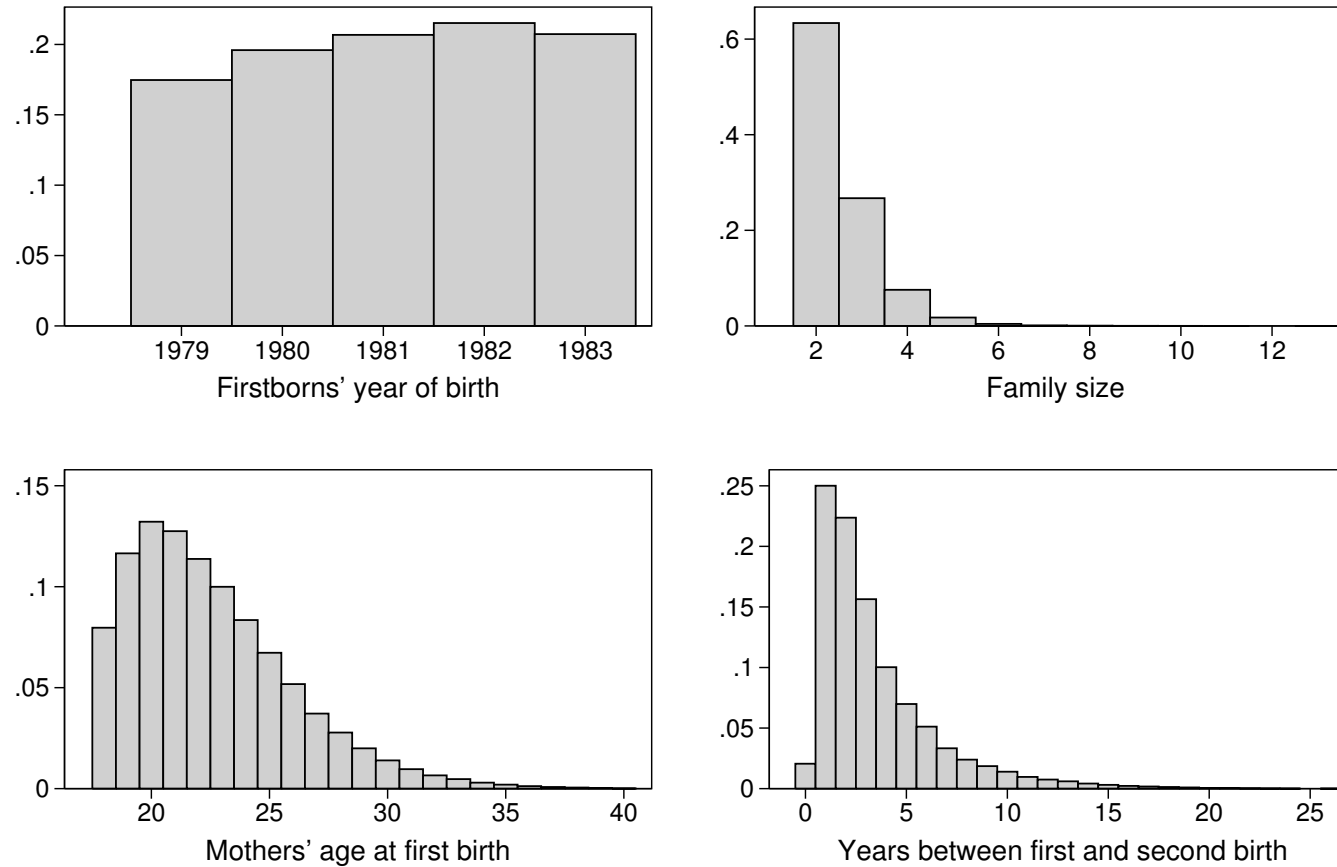
*Notes* : The table summarizes empirical evidence on the effect of family size on labor market outcomes of the firstborn child, using multiple second births as an instrument. Sample means, coefficients, and standard errors are taken from the original articles, and the confidence intervals are calculated based on the estimates. Sample means correspond to: <sup>a</sup> firstborn children; <sup>b</sup> all singleton births; <sup>c</sup> dummy variable that equals one if the child's highest completed grade is lower than the mode of highest grade completed by age in years, quarter of birth and state, and zero otherwise; <sup>d</sup> children per family, excluding twins; <sup>e</sup> indicator variable running from zero (excellent health) to -12 (lowest health status) based on examinations of individuals to be allowed to do military training.

Figure 1: Share of Twins over Time in Austria and Birth Cohorts of Second Births in Sample



Notes : The upper first graph shows the share of twins at parity one and two over time in Austria. The second graph shows the cumulative distribution function for the year of second birth for families used to estimate firstborns' human capital outcomes (N=107,270).

Figure 2: Distributions of Key Variables



Notes : These figures are based on the full sample, which is used to estimate firstborns' human capital outcomes ( $N=107,270$ ) and show the distribution of firstborn's year of birth (top left), completed family size (top right), mothers' age at first birth (bottom left), and years between first and second birth (top right).

**Table 2: First Stage Relationships**

	(1)	(2)	(3)
twin 2	0.777 [0.740,0.815]		
same sex		0.061 [0.052,0.070]	
girl 1,2			0.079 [0.066,0.092]
boy 1,2			0.043 [0.031,0.056]
No. of observations	107,270	106,346	106,346
Mean of dependent var.	2.5	2.5	2.5
F-statistics on weak IV(s) <sup>a</sup>	1,004.4 <sup>a</sup> / 1,673.2 <sup>b</sup>	178.9 <sup>a</sup> / 178.8 <sup>b</sup>	96.7 <sup>c</sup>

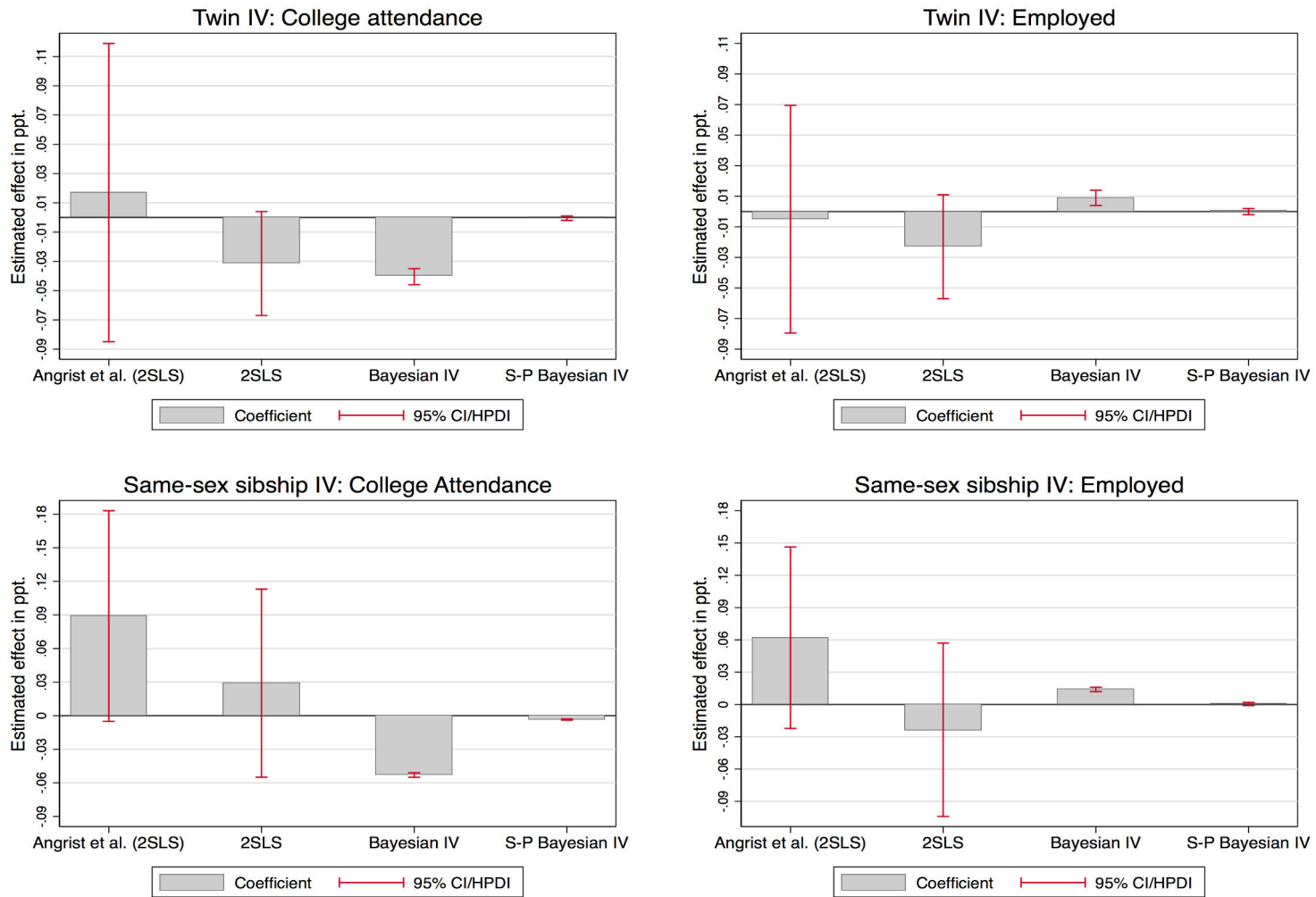
*Notes* : This table summarizes the least square estimation of the effect of (1) a twin birth at second birth, (2) a same-sex sibship, and (3) sex-specific same-sex sibship on completed family size (i.e. number of children). Each specification controls for the firstborn's year of birth and mother's age at first and second birth. We consider a 95 percent confidence interval, based on robust standard errors, allowing for heteroskedasticity of unknown form in the brackets below: <sup>a</sup> Cragg-Donald F statistic; <sup>b</sup> Kleibergen-Paap Wald F statistic; and the <sup>c</sup> Angrist-Pischke multivariate F Test of excluded IVs. The different number of observations for the twin and the same-sex sibship IV occur because twin births are excluded from the same sex specifications.

**Table 3: The Effect of Family Size on Firstborns' Human Capital Outcomes**

	(1)	(2)	(3)	TWIN IV			SAME-SEX SIBSHIP IV		
				(4)	(5)	(6)	(7)	(8)	(9)
	Mean	OLS	Bayesian	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV
College attendance	0.29	-0.012 [-0.015,-0.009]	-0.015 [-0.017,-0.012]	-0.031 [-0.067,0.004]	-0.04 [-0.046,-0.035]	0.0002 [-0.002,0.001]	0.029 [-0.055,0.113]	-0.053 [-0.055,-0.051]	-0.0035 [-0.004,-0.003]
Employment	0.80	-0.017 [-0.021,-0.014]	-0.015 [-0.017,-0.014]	-0.023 [-0.057,0.011]	0.009 [0.004,0.014]	0.0009 [-0.002,0.002]	-0.024 [-0.104,0.057]	0.014 [0.012,0.016]	0.0005 [-0.001,0.002]
White-collar	0.71	-0.033 [-0.038,-0.029]	-0.034 [-0.037,-0.032]	0.003 [-0.038,0.045]	-0.029 [-0.036,-0.023]	-0.0008 [-0.003,0.001]	0.093 [0.001,0.185]	-0.031 [-0.033,-0.028]	-0.0007 [-0.001,-0.001]
Wage	83.71	-1.675 [-1.993,-1.358]	-1.827 [-2.02,-1.63]	2.765 [-0.517,6.047]	-1.45 [-1.98,-0.93]	2.72 [-6.74,11.79]	3.736 [-3.814,11.285]	-2.08 [-2.29,-1.88]	-3.34 [-10.20,3.04]

*Notes* : This table summarizes the effects of family size on firstborns' educational and labor market outcomes. The means of dependent variables are shown in column (1), columns (2) and (3) show the OLS estimates and Bayesian linear regression results. Columns (4) to (6) show the estimation results for the twin IV using 2SLS, a Standard Bayesian IV, and a semi-parametric Bayesian IV approach. Columns (7) to (9) show the same for the same-sex sibship IV. Each cell represents the results from a separate regression controlling for a firstborn's year of birth and mother's age at first and second birth. The 95 percent confidence intervals based on robust standard errors, allowing for heteroskedasticity of unknown form, and the 95 percent HPD intervals for the Bayesian results are shown in brackets. The number of observations for college attendance and employment are 107,270 in columns (1) to (6) and 106,346 in columns (7) to (9). The white-collar indicator and wage is only available for a subsample of individuals. Here, the respective number of observations is 77,488 (columns (1)-(6)) and 76,836 (columns (7)-(9)) for white-collar and 83,303 (columns (1)-(6)) and 82,606 (columns (7)-(9)) for wage regressions.

Figure 3: Graphical Comparisons of Different IV-Estimators



Notes : The bar graphs summarize our estimation results for two outcomes based on two different IVs, along with comparable estimates from Angrist *et al.* (2010). The top, left panel summarizes the estimates for college attendance using the twin IV, the top right panel summarizes the estimates for employment using the twin IV, the bottom left panel summarizes the estimates for college attendance using the same-sex sibship IV, and the bottom right panel summarizes the estimates for employment using the same-sex sibship IV. In each graph, the first bar displays the results obtained from Angrist *et al.* (2010), the second bar captures our 2SLS estimate, the third bar captures our standard Bayesian IV estimate, and the fourth bar shows our semi-parametric Bayesian IV estimate. The whiskers in each bar show the 95 percent confidence intervals or the highest posterior density (HPD) intervals, in the case of Bayesian estimate.

**Table 4: The Effect of Family Size on Firstborns' Health Outcomes**

	(1)	(2)	(3)	TWIN IV			SAME-SEX SIBSHIP IV		
				(4)	(5)	(6)	(7)	(8)	(9)
	Mean	OLS	Bayesian	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV
Total expend.	583.88	-1.802	2.52	-75.435	-0.74	15.72	136.423	2.26	69.93
		[-35.315,31.711]	[-3.50,8.57]	[-226.133,75.264]	[-19.02,18.94]	[-19.43,51.74]	[-607.917,880.762]	[-14.41,19.88]	[-235.58,329.32]
Med. attendance	218.56	1.365	0.45	4.152	2.41	11.28	91.952	4.70	-26.50
		[-3.226,5.957]	[-0.51,1.41]	[-41.082,49.386]	[-7.75,13.11]	[-15.94,37.55]	[-26.640,210.544]	[-0.45,9.56]	[-219.81,137.72]
Drug use	86.60	-9.691	-4.45	-16.328	-3.60	2.02	-74.039	-8.98	-0.28
		[-19.608,0.227]	[-6.67,-2.21]	[-71.790,39.134]	[-20.56,13.34]	[-1.78,5.86]	[-354.862,206.784]	[-19.92,1.73]	[-72.49,60.88]
Days in hospital	1.11	0.034	0.044	-0.125	-0.008	0.00002	-0.301	0.035	-0.024
		[-0.094,0.162]	[0.015,0.068]	[-0.647,0.396]	[-0.29,0.28]	[-0.01,0.011]	[-3.062,2.460]	[-0.09,0.17]	[-0.10,0.03]

*Notes* : This table summarizes the effects of family size on firstborns' health outcomes. The means of the dependent variables are shown in column (1), and columns (2) and (3) show the OLS estimates and Bayesian linear regression results. Columns (4) to (6) show the estimation results for the twin IV using 2SLS, a Standard Bayesian IV, and a semi-parametric Bayesian IV approach. Columns (7) to (9) show the same for the same-sex sibship IV. Each cell represents results from a separate regression controlling for firstborn's year of birth and mother's age at first and second birth. The 95 percent confidence intervals based on robust standard errors allowing for heteroskedasticity of unknown form and the 95 percent HPD intervals for Bayesian results are shown in brackets. The number of observations is 17,746 in columns (1) to (6) and 17,577 in columns (7) to (9).

**Table 5: The Effect of Family Size on Firstborns' Human Capital Outcomes by Mothers' Educational Attainment**

	(1)	(2)	(3)	TWIN IV			SAME-SEX SIBSHIP IV		
				(4)	(5)	(6)	(7)	(8)	(9)
	Mean	OLS	Bayesian	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV	2SLS	Standard Bayesian IV	Semi-parametric Bayesian IV
<b>Panel A: Compulsory schooling only</b>									
College attendance	0.11	-0.009 [-0.014,-0.005]	-0.009 [-0.012,-0.006]	-0.000 [-0.054,0.053]	-0.013 [-0.021,-0.008]	0.0004 [-0.001,0.003]	0.028 [-0.111,0.168]	-0.018 [-0.020,-0.015]	-0.0004 [-0.003,0.001]
Employment	0.78	-0.021 [-0.027,-0.014]	-0.020 [-0.024,-0.017]	-0.118 [-0.198,-0.038]	0.024 [0.011,0.040]	0.001 [-0.004,0.004]	-0.055 [-0.244,0.134]	0.037 [0.033,0.040]	-0.005 [-0.008,-0.001]
White-collar	0.56	-0.039 [-0.048,-0.030]	-0.022 [-0.027,-0.016]	0.047 [-0.052,0.146]	0.019 [-0.006,0.028]	-0.0002 [-0.005,0.005]	0.180 [-0.068,0.427]	0.030 [0.024,0.035]	0.001 [-0.005,0.005]
Wage	78.78	-1.344 [-1.923,-0.765]	-0.165 [-0.995,0.166]	6.494 [0.371,12.617]	3.726 [2.748,4.695]	0.294 [-15.688,15.748]	0.757 [-14.725,16.239]	4.333 [4.007,4.662]	-1.106 [-13.721,14.813]
<b>Panel B: Graduates from low track</b>									
College attendance	0.25	-0.019 [-0.024,-0.014]	-0.012 [-0.015,-0.008]	-0.035 [-0.086,0.016]	-0.031 [-0.039,-0.022]	-0.005 [-0.011,-0.0002]	-0.004 [-0.141,0.134]	-0.034 [-0.037,-0.031]	-0.002 [-0.003,-0.0002]
Employment	0.80	-0.014 [-0.018,-0.009]	-0.018 [-0.021,-0.015]	0.028 [-0.019,0.075]	0.035 [0.027,0.043]	0.0002 [-0.004,0.003]	-0.044 [-0.175,0.088]	0.047 [0.044,0.050]	0.000 [-0.003,0.002]
White-collar	0.72	-0.034 [-0.040,-0.027]	-0.014 [-0.018,-0.009]	-0.030 [-0.091,0.030]	0.020 [0.009,0.031]	-0.000 [-0.003,0.003]	-0.038 [-0.181,0.104]	0.039 [0.035,0.043]	-0.000 [-0.003,0.003]
Wage	82.90	-1.453 [-1.923,-0.983]	0.234 [-0.064,0.534]	-1.500 [-6.047,3.047]	3.513 [2.723,4.348]	3.958 [-11.018,16.515]	3.492 [-7.980,14.963]	5.746 [5.460,6.031]	-2.941 [-15.843,14.375]
<b>Panel C: Graduates from high track</b>									
college attendance	0.64	0.003 [-0.008,0.014]	0.011 [0.004,0.018]	-0.012 [-0.130,0.107]	0.018 [-0.0001,0.037]	-0.001 [-0.012,0.004]	0.010 [-0.244,0.264]	0.024 [0.017,0.031]	0.000 [-0.005,0.006]
Employment	0.79	-0.017 [-0.027,-0.007]	-0.014 [-0.021,-0.008]	-0.055 [-0.166,0.055]	0.043 [0.026,0.060]	0.000 [-0.005,0.005]	-0.048 [-0.292,0.197]	0.054 [0.048,0.060]	0.000 [-0.002,0.006]
White-collar	0.90	-0.009 [-0.019,-0.000]	-0.002 [-0.009,0.004]	0.017 [-0.074,0.108]	0.035 [0.017,0.052]	-0.007 [-0.017,0.005]	0.089 [-0.030,0.209]	0.065 [0.059,0.072]	0.000 [-0.003,0.006]
Wage	86.81	-1.527 [-2.648,-0.406]	2.564 [1.816,3.314]	5.088 [-6.516,16.692]	7.314 [5.304,9.176]	5.007 [-31.737,36.338]	14.163 [-8.783,37.109]	9.345 [8.615,10.083]	0.075 [-21.705,23.239]

This table summarizes the effects of family size on firstborns' educational and labor market outcomes by mothers' educational attainment. The means of the dependent variables are in column (1), columns (2) and (3) show the OLS estimates and Bayesian linear regression results. Columns (4) to (6) show the estimation results for the twin IV using 2SLS, a Standard Bayesian IV, and a semi-parametric Bayesian IV approach, and columns (7) to (9) show the same for the same-sex sibship IV. Each cell represents results from a separate regression controlling for firstborn's year of birth and mother's age at first and second birth. The 95 percent confidence intervals based on robust standard errors allowing for heteroskedasticity of unknown form and 95 percent HPD intervals for Bayesian results are in brackets below.



**Table 6: Number of Components Required by the Semi-Parametric Bayesian IV**

	TWIN IV			SAME-SEX SIBSHIP IV		
	Lower bound	Median	Upper bound	Lower bound	Median	Upper bound
College attendance	5	5	6	5	5	6
Employment	4	5	5	4	4	6
White-collar	4	5	6	4	4	4
Wage	7	7	8	7	8	8
Total expenditures	18	21	24	17	20	23
Medical attendance	11	11	12	10	10	11
Drug use	18	21	25	12	13	21
Days in hospital	11	13	14	11	12	14

*Notes* : The table summarizes the distribution of the number of components required by the semi-parametric Bayesian IV model. The first and fourth column correspond to the lower bound of the corresponding 95 percent HPD interval, the second and fifth column to the median, and the third and sixth columns to the upper bound of the HPD interval for the twin IV and same-sex sibship IV, respectively.