

IZA DP No. 2437

## Intergenerational Transmission of Fertility Patterns in Britain

Alison L. Booth  
Hiau Joo Kee

November 2006

# Intergenerational Transmission of Fertility Patterns in Britain

**Alison L. Booth**

*Australian National University, University of Essex,  
CEPR and IZA Bonn*

**Hiau Joo Kee**

*Australian National University*

Discussion Paper No. 2437  
November 2006

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Intergenerational Transmission of Fertility Patterns in Britain<sup>\*</sup>

Recent studies by economists exploring the nexus between culture and fertility have focused on cultural transmission from the origin country rather than the origin family. Our paper extends this avenue of research by investigating how *family-specific* 'cultural transmission' can affect fertility rates. In this context, we define 'culture' as referring to intra-family norms, and 'cultural transmission' refers to the transfer of these norms across generations within a family. We also allow for peer-group influences through the inclusion of controls for age cohorts and for non-English speaking country of birth. Following the methodology of Miranda (2005) and Machado and Santos Silva (2005), we estimate count data quantile regression models. Using unique data from the British Household Panel Survey, we find that a woman's origin-family size is positively associated with her own completed fertility in the destination family and that her country of birth also matters. The effect of origin family size increases as we move from the lower to the upper tail of the conditional fertility distribution. For a sub-sample of continuously partnered men and women, *both* partners' origin-family sizes significantly affect destination-family fertility. Our findings are robust to a number of specification checks.

JEL Classification: F22, J15, J16, Z10

Keywords: completed fertility, origin family size, inter-generational effects, counts quantile regression

Corresponding author:

Alison L. Booth  
Economics Department  
University of Essex  
Wivenhoe Park CO4 3SQ  
Essex  
United Kingdom  
E-mail: [albooth@essex.ac.uk](mailto:albooth@essex.ac.uk)

---

<sup>\*</sup> We are grateful to Wiji Arulampalam, John Ermisch, Alfonso Miranda, Joao Santos Silva, and seminar participants at the Australian National University for helpful suggestions. Special thanks are also due to Alfonso Miranda for providing us with his code. The data were made available through the UK Data Archive. The data were originally collected by the ESRC Research Centre on Micro-social Change at the University of Essex, now incorporated within the Institute for Social and Economic Research. Neither the original collectors of the data nor the Archive bear any responsibility for the analyses or interpretations presented here. HJ Kee would like to thank the University of Essex for its hospitality during the writing of this paper. Part of the research was also funded by an ARC Discovery Project Grant.

## I. INTRODUCTION

In this paper we use the British Household Panel Survey to study inter-generational fertility patterns. We investigate the degree to which origin-family characteristics affects women's subsequent completed fertility in their destination families. In particular we are interested in how, within any one biological family, the fertility of one generation affects the fertility of the next. We also explore how intergenerational fertility behaviour varies with birth order status in the family of origin.

Demographers and sociologists have, for over a hundred years, been interested in intergenerational fertility patterns.<sup>2</sup> Genetic differences, either in the desire to have children or in the ability to have them, were initially stressed.<sup>3</sup> Subsequent studies emphasised the possibility of intergenerational transmission of contraceptive technologies and know-how, while others argued that this information might instead be transmitted from the relevant peer group rather than from mother to daughter. Other studies highlighted the role that norms in the origin family might play in affecting children's subsequent family planning decisions. Fertility norms of other reference groups – based on friendship, ethnicity, social class and religion - have also been emphasised (see for example Westoff and Potvin, 1967). Following Bisin and Verdier (2000, 2001), we label these norms as 'culture' and the transmission of these norms across generations as 'cultural transmission'. These influences are our primary interest in this paper.

Since economists typically assume that preferences are beyond the scope of their analysis, it is perhaps unsurprising that there has been relatively little work by economists on cultural transmission. Instead, economists focused on other determinants of fertility, especially the

---

<sup>2</sup> For example see Pearson and Lee (1899).

<sup>3</sup> More recently, genetic influences are again being explored in studies such as Rodgers et al (2001), Miller et al (1999) and Kohler et al (1999). When fertility norms and birth control technology do little to constrain individuals' fertility choices, genetics may play an important role in fertility outcomes (Kohler *et al* 1999). While our data source does not allow us to explore this genetic component, it does provide important new information about origin-family characteristics including family size.

relationships between fertility and investments in human capital, and between fertility and life expectancy.<sup>4</sup> However, recent studies argue that cultural transmission plays an important role in explaining certain economic phenomena. For example, influential papers by Borjas (1992, 1995) and Bisin and Verdier (2000, 2001) invoke cultural transmission as an important explanation of the non-assimilation of immigrants. Blau (1992), Guinnane et al. (2002) and Fernandez and Fogli (2005) explore how cultural transmission can explain heterogeneity in US fertility rates across immigrant groups.<sup>5</sup> The focus of these studies is on modelling culture as in the country of origin rather than the family of origin.<sup>6</sup>

Our paper extends this new avenue of economic research by investigating how *family-specific* cultural transmission can affect fertility rates. Demographers and sociologists have already documented the importance of this avenue of transmission, but we use a more appropriate econometric methodology. We also use new data containing a rich set of origin-family controls as well as information on completed fertility in the destination-family. In our analysis we define ‘culture’ as intra-family norms, and ‘cultural transmission’ refers to the transfer of these norms across generations within a family. We also allow for peer-group influences through the inclusion of controls for age cohorts and for country of birth. Our analysis is conducted at the family level.

It has sometimes been suggested that different birth order children may systematically differ in their receptiveness to transmission of family culture (see for example Hendershot (1969), Ejrnaes and Portner (2004) and Black *et al* (2005)). Hendershot (1969) suggests there are differences in the degree of ‘socialization’ across birth order and that first born are more likely to conform to parental norms as they are more susceptible to social pressure from parents than

---

<sup>4</sup> For examples of the former, see Becker (1960), Becker and Barro, (1988), Becker and Lewis, (1973), Becker and Tomes, (1976), Hanushek (1992) and Ermisch (2003). For the latter, see Lee (2003) and Livi-Bacci (2001).

<sup>5</sup> Fernandez and Fogli (2004) also model the effects on labour force participation.

<sup>6</sup> For recent research by demographers on inter-generational transmission of fertility patterns, see Murphy and Wang (2001), Murphy and Knudsen (2002) and Axinn et al. (1994).

subsequent children. More recently Sullo way (1996) finds empirical support for the hypothesis that first-borns are more conformist than later-borns.<sup>7</sup> For this reason, we also control for birth order in our analysis.

Women typically do not consider fertility decisions in isolation. They have partners whose preferences need to be accommodated. Each partner in the destination family comes from two separate origin families, which may be characterized by different family norms and fertility histories. A strength of the British Household Panel Survey is that it provides information for both partners and we can therefore control for their different family backgrounds. Of course intra-household bargaining power might affect the relative weighting given to each partner's fertility norms.<sup>8</sup> We use some proxies that might pick up differences in bargaining power (such as age difference and educational differences between partners).

A challenge for studies attempting to measure the impact of culture on fertility is that there may be some unobserved characteristic that is correlated with both the cultural proxy and with the independent variable, completed fertility. For example, highly educated parents in the origin family are less likely to have had large families, and highly educated parents are also more likely to have higher educated children. Hence any estimated correlation between destination family size and origin family size might work through omitted human capital effects rather than through cultural transmission.

Fortunately our data set has a rich sets of controls for both origin-family background and for destination-family human capital attainment. Of course human capital in the destination family may be partly endogenous. Young women who anticipate having a large family in the future may be less likely to invest in human capital as young adults, since they will be less likely to

---

<sup>7</sup> Sullo way (1996) explains this finding by sibling competition for parental attention and care. First-borns typically find conforming to the parental model to be a successful strategy. Since later-born children cannot displace the first-born from that niche, they adopt an alternative strategy to win parental attention.

<sup>8</sup> Details see Lundberg and Pollak, 1996.

gain the returns though continuous labour market participation. We therefore estimate some specifications of completed fertility in which we control only for the observed characteristics of the origin family, including parental human capital, demographics, family size and birth order.<sup>9</sup> We then estimate expanded specifications in which we also include as a control the highest educational level of the woman in the destination family. When we compare the estimated coefficients to family size and birth order in the parsimonious and expanded specifications, we find they are remarkably similar. This suggests that our estimated coefficients to our variables of interest – origin family size and birth order – are not suffering from omitted variable bias.

As an extension to the existing literature, in this paper we model completed fertility counts utilising a quantile regression (QR) approach, which is a relatively less restrictive than alternative methods. The QR framework allows for the characteristics to have different returns at different quantiles. The effects of covariates on the location, scale and shape of the conditional fertility distribution can be easily estimated using this framework. This is a major advantage compared to the alternatives.

The remainder of our paper is set out as follows. Section II describes the data source and the explanatory variables. Section III presents estimates of completed fertility using quantile regression techniques on jittered counts data, following the procedure used by Miranda (2005) and Machado and Santos Silva (2005). Section IV presents estimates of the determinants of children's perceptions of the importance of having children. The final section concludes.

## **II. THE DATA AND VARIABLES**

### **II.1 The Data Source**

---

<sup>9</sup> These family background factors have a strong effect on individuals' educational attainment, as shown in papers by, inter alia, Black et al. (2005), Conley (2000), and - using the BHPS - Booth and Kee (2005).

The data source is the British Household Panel Survey (BHPS). The BHPS is a nationally representative random-sample survey of private households in Britain. The same individuals are re-interviewed in successive waves and, if they split off from original households, all adult members of their new households are also interviewed. Children are interviewed separately once they reach the age of 16. Although limited information on family background was collected in earlier waves, the questionnaire was expanded in the 13<sup>th</sup> wave to elicit additional information about family and parental background, and the childhood home. Of particular interest are the new variables about sibling numbers and birth order. We use these to investigate the degree to which family size and birth order within the origin family affect women's subsequent fertility decisions.<sup>10</sup> Other family background variables allow us to control for family-level heterogeneity. We also utilise information from earlier waves to construct a measure of completed fertility for our estimating sample of 2103 women aged 45 to 65 years in 2003.

Our main estimating sample consists of 2103 women aged between 45 and 65 years, and with valid information on the main variables (number of biological children, family size and birth order). We excluded from the sample individuals aged less than 45 in order to ensure that respondents had completed their fertility. For each woman in this sample, we have the number of siblings in her *origin family* as well as the number of biological children she has produced in the *destination family*. We also have a number of attributes about the origin family. These are clearly exogenous to her subsequent completed fertility, since they are measured typically when the individual was aged 14 or less. In addition, we also undertake some analysis of two further sub-samples, as will be described in more detail in Section III below. The first of these is a sub-sample of continuously married women and their husbands. The second is a sub-sample of

---

<sup>10</sup> These variables are retrospective and with retrospective data there are always issues about potential recall error. However, the variables in which we are interested relate to attributes that are unlikely to be forgotten; it is hard to imagine that anyone within our sample of interest – 28-55 year olds – would be likely to forget the number of siblings or their own birth order.



children from the destination family who were also surveyed about their perceptions of the importance of having children.

## II.2 The Variables

### *Dependent Variable: Completed Fertility in the Destination Family*

Since the BHPS does not provide a derived variable for completed fertility, we constructed it from responses across earlier waves as described in detail in the Data Appendix A.1.<sup>11</sup> Figure 1 and Table 1 give the distribution of the number of biological children in the destination family. The majority of women in our estimating sample of 2103 women aged 45-65 have had around 2 to 3 children. The sample mean is 2.27 and the variance is 1.73, which suggests that the data exhibit under-dispersion in raw terms.<sup>12</sup>

The first column of Table 1 gives the number of observations for each count while the second gives the percentages. Thus there are 269 women (12.8%) with only one child, as compared with 846 women (40.2%) with two children, 496 (23.6%) with three children, and 193 (9.2%) with four children.

The last column of Table 1 reports the predicted counts of completed fertility from a Poisson count data model in which only the constant is included (see Winkelmann, 2003 for a full exposition of count data models). This reveals that there is under-prediction of the raw data for the count of 2 children (the actual is 40.23% while the predicted from the Poisson model is only 26.6%) and also for 3 children. Indeed, a goodness-of-fit  $\chi^2$  statistic test shows  $\chi^2=322.06$ , which leads to the rejection of the null hypothesis that the data follow a Poisson distribution.

---

<sup>11</sup> Large families were less common in Britain than in less developed countries, and the maximum value for completed fertility in our sample is 12 children.

<sup>12</sup> Over (under)-dispersion refers to the situation where variance is greater (smaller) than the mean. For details see Winkelmann (2003).

Our basic cultural transmission hypothesis is that individuals' fertility preferences are affected by their origin family size. But with our data we actually observe *completed* fertility rather than preferences for children. Some women in our sample may have wanted children and yet have been unable to have them, and these women will be in the zeroes together with women who chose zero. But notice that in our data there is a relatively low percentage with no biological children, at 9.5%. This raises another unique feature of fertility data which makes the econometric modelling challenging. Besides the typical under-dispersion, the observed zeroes have multiple meanings and the fertility decision is a sequential process. Parents decide whether or not to have any children; and conditional on the current number of children, whether or not to have new child. As pointed out in Miranda (2003), reasons that lead a couple to have their first child may be different from the reasons that may lead them to have further children. In Section III we discuss the choice of econometric model.

***Other Variables: Origin Family Size and Birth Order***

Respondents in wave 13 were asked some additional questions about origin-family size and their own birth order within that family, as well as family background attributes. The question about family size (question D108) was: "How many brothers and sisters have you ever had?" This was immediately followed by the question: "So including yourself, there were (D108+1) children in your family?" We used this information to construct a variable for the total number of children in the origin family, which is top-coded at 10 children. The next question asked "Where were you born in relation to your brother(s) and sister(s), that is, were you the first, second, third or subsequent child?" There followed a list of up to 10 possibilities, with the 10<sup>th</sup> top-coded as "tenth (or later)."<sup>13</sup> As detailed in the Data Appendix A.2, we use a

---

<sup>13</sup> Unfortunately the BHPS does not provide information about the gaps between siblings.

transformation of this birth order information as a control in our estimation of the correlates of completed fertility.<sup>14</sup>

Table 2 presents cross-tabulations of completed fertility for women aged 45 to 65 by their origin-family size. It also reports the number of observations in each origin-family size group. Inspection of the table reveals a positive correlation between origin and destination family size. For example, the second row shows that 92 % of 502 women from 2-child origin-families have 3 or fewer children, while the seventh row shows that 73 % of 88 women from 7-child origin-families have 3 or fewer children.

### ***Other Sources of Fertility Norms***

As noted in the Introduction, fertility decisions may be affected not only by parental norms but also by fertility norms characterizing other reference groups. These groups might be based on friendship, ethnicity, social class and religion. Bisin and Verdier (2001) formally model intergenerational cultural transmission involving either vertical socialisation to the parents' trait or horizontal socialisation to the dominant trait of a group in the population. They explore population dynamics when these socialisation avenues are either substitutes or complements. Owing to data limitations we are unable to test this hypothesis directly with our data. However we do control for some of these factors by adding dummy variables for age cohorts and whether or not the respondent comes from a Non-English-Speaking-Background (NESB), and so we are able to estimate the relative impacts of origin family size and background, as well as external factors such as birth cohorts. We do not have enough observations to use more disaggregated measures of ethnicity, since in our sample around 97% of the respondents were born in the UK.

---

<sup>14</sup> See Booth and Kee (2005) for further details. The fact that cell sizes across birth order categories are relatively small for some of our larger families suggests that it is important to find a parsimonious way of representing the data. To this end, we converted responses to the birth order question into an index, which parsimoniously represents the data and reduces almost to zero the correlation between family size and birth order.

We are also missing some variables that we would like to include. For example, siblings' fertility might also matter but we do not have information about this, nor do we have data on age gaps between siblings and the sex of siblings in the survey.<sup>15</sup> We also do not have appropriate religion measures which might affect fertility.<sup>16</sup>

Figure 2 shows how the mean and variance of completed fertility vary across each year of age (that is, from 45 to 65). The mean is higher for older women as expected. As noted above, we include age cohort dummies in our estimation to control for peer effects and for contraceptive technology changes (since the older cohorts would not have had access to the oral contraceptive). Interestingly, the variance in completed fertility is also slightly higher for older women.

### ***Other Sources of Heterogeneity***

Table 3 gives the means of the variables used in our analysis, together with a brief definition of each. For our main estimating sample of 2,103 women, 25.4% are between the ages of 46 and 50, 26.6% are between the ages of 51 and 55, 27.5% are between the ages of 56 and 60, and 20.3% are between the ages of 61 and 65 years old. The mean origin-family size (including the respondent herself) is 3.757, and about 32.5% women have no qualification or an undefined

---

<sup>15</sup> There are two possible ways in BHPS to find out the sex composition of family members: One can either work it out from the Household relationship grid; or to find it out from the special fertility modules such as *xCHILDNT* and *xCHILD* that are available in only specific waves. The first option is not feasible for this study. This is because some children do not stay with their parents due to marriage break ups or other family living arrangements; and some respondents have mature children at the time of the interview. The second option also has its drawbacks. The special fertility modules *xCHILDNT* were only asked in wave 2, 11 and 12. Although all respondents were asked in the second waves, only new respondents from the Scotland or Wales were asked in wave 11 and 12. Another fertility module *xCHILD* was only available in wave 12 and 13. Similarly, although all respondents were asked in wave 12, only respondents from Northern Ireland and those who have at least one or more children under age 16 were interviewed. Hence we could by this route only work out the children's gender composition for some of the respondents, and this would not necessarily be representative of all respondents.

<sup>16</sup> Wave 13 only provides respondents' religion denomination at 2003. This is likely to suffer from potential endogeneity. In addition, wave 13 grouped all non-Christian religions into one, which makes it impossible disaggregate the effect of different religions such as Muslim and Buddhist. We therefore decided not to include religion in our analysis.

qualification. About 74.2% women in the sub-sample are currently legally married, and the mean age of first marriage of all women is 23.96.

Wave 13 of the BHPS also provides unique information about origin family attributes that allows us to control for family-specific heterogeneity. Table 3 in the Appendix also gives the means of these variables. The presence of books in the parental home when the mother was a child forms a proxy for family-specific attitudes to education and the like. These may affect subsequent fertility decisions either directly, or indirectly through their correlation with subsequent educational attainment (Booth and Kee, 2005). Households with many books are likely to have a more positive attitude to learning through the written word than are households with few or no books.<sup>17</sup> We proxy parental wealth in the origin family by dummy variables taking the value one if the mother had at least high school or qualification, and zero otherwise, and likewise for the father. We also use a dummy variable indicating whether or not the mother worked when the respondent was aged 14. Area-specific factors are captured by a set of variables indicating the type of area in which the family mostly lived when the respondent was a child.<sup>18</sup> This could proxy area-specific norms. To proxy the effect of broken families on respondent fertility patterns, we include a dummy variable showing whether or not respondent lived with both biological parents until the age 16.

Other variables that are likely to affect women's fertility in their destination family include the length of their period of fecundity and their educational attainment. The former can be proxied by age at first marriage. If women are highly educated they are likely delay marriage, or if they marry late regardless of education, hence they will have shorter period over which to

---

<sup>17</sup> Respondents were asked: "Thinking about the time from when you were a baby until the age of ten, which of the following statements best describes your family home: There were a lot of books in the house; There were quite a few books in the house; There were not very many books in the house; Don't know." We constructed dummy variables for "a lot of books in the house" and "quite a few books in the house". The base in the regressions is "not many books in the house".

<sup>18</sup> The precise question about area of residence was: "Please look at this card and tell me which best describes the type of area you mostly lived in from when you were a baby to 15 years." Responses are described in Appendix Table 3. The base for the regressions is "move around when during childhood".

reproduce. We control for these factors with age of marriage and with predetermined variables affecting educational attainment. In our robustness checks in Section III.3, we also experiment with introducing explicit measures for educational attainment.

### **III. THE ESTIMATES**

#### **III.1 The Econometric Model**

Unlike most types of microeconomic data, completed fertility data are generally under-dispersed. In this situation, the equi-dispersion assumption of the standard Poisson count model is violated and inference based on the estimated standard errors is no longer valid (for details see Cameron and Trivedi, 1986, and Winkelmann and Zimmermann, 1994). This uncommon feature of fertility data has led to the development of more flexible statistical methods based on different generalisations of the Poisson distribution. Examples include a generalised Poisson regression of Wang and Famoye (1997); an inflated Gamma count model of Melkersoon and Rooth (2000); and a Gamma count model accounting for nonexponential waiting times between events of Winkelmann (1995).

The assumption behind these fully parametric count models is that the shape of the outcome distribution remains unchanged even when the value of explanatory variables changes. In other words, explanatory variables have a homogenous effect at all regions of the conditional outcome distribution (Winkelmann, 2003). In spite of this, it has been well documented that the fertility decision is derived by two different decision processes.<sup>19</sup> To take this into consideration, a hurdle type probabilistic model has been widely adopted in previous studies.<sup>20</sup>

---

<sup>19</sup> In the first stage, parents decide whether or not to have any children; and after the first stage decision is reached, parents then consider whether or not to have any further child conditional on the first stage. Because the two tails of the distribution are generated by two different processes, it is important to assess the effect of the covariates on different regions of the distribution (Miranda, 2003).

<sup>20</sup> Examples include the three-stage hurdle of Miranda (2003) and the modified generalised Poisson hurdle of Santos Silva and Covas (2000).

Instead of using a heavily parameterised model to analyse how the effects of covariates vary across different regions of the conditional distribution of interest, Machado and Santos Silva (2005) proposed the use of Quantile Regression (QR) on count data. Since the data are naturally discrete, some smoothness must be artificially imposed on the problem. They showed that it is possible to smooth the data in a way that allows inference to be performed using standard quantile regression techniques.<sup>21</sup> This not only enables the researcher to study the impact of the regressors on each quantile of the distribution, but it also requires less distributional and probabilistic assumptions. For this reason we will use a QR approach similar to that of Miranda (2005), who also utilise QR in the context of fertility.

Consider random variables  $y_i$  and  $x_i$ , in which  $y_i$  is a discrete and nonnegative count variable given regressors  $x_i$ , and the subscript  $i$  represents realisations of random variables of the  $i$ -th sample observations. Machado and Santos Silva (2005) suggested constructing a continuous and random auxiliary variable  $z_i = y_i + u_i$ , whose quantiles have a one-to-one relation with  $y_i$ , and  $u_i$  is a random variable independent of  $y_i$  and  $x_i$  uniformly distributed in the interval  $[0,1)$ . This approach utilised a specific form of jittering introduced by Stevens (1950).

In the content of QR it is necessary to specify the form of  $Q_z(\alpha | x)$  and the associated transformation  $T(z_i, \alpha)$ . A parametric representation of  $Q_z(\alpha | x)$  can be specified as

$$Q_z(\alpha | x) = \alpha + \exp(x'\gamma(\alpha))$$

Moreover,  $\gamma(\alpha)$  can be estimated by running a linear QR of

$$T(z, \alpha) = \begin{cases} \log(z - \alpha) & \text{if } Z > \alpha \\ \log(\xi) & \text{if } Z \leq \alpha \end{cases}$$

---

<sup>21</sup> QR on count data is complicated by the combination of a non-differentiable objective function with a discrete dependent variable. Huber (1981) pointed out that the conditional quantiles may not be approximated by a Taylor expansion under such condition. Hence the application of QR to count data is only possible if some smoothness is artificially integrated into the problem. In this study we follow the Machado and Santos Silva (2005) approach. However there are other alternatives; see for example Lee (1992) and Efron (1992).

on  $x$  with  $\xi$  represents a suitably small positive number. The derived vector of parameter  $\gamma(\alpha)$  are estimated by means of a standard QR of  $T(z_i, \alpha)$  on the vector of explanatory variables  $x_i$ . For details of the asymptotic properties, covariance matrix and the mathematical derivations, refer to Machado and Santos Silva (2005).

Recall that random variable  $u_i$  was introduced to the model by technical reasons. To minimise the uniform noise, Machado and Santos Silva (2005) consider the process of “average-jittering”. This method averages the QR estimates for  $m$  “jittered” samples  $\{y_i + u_i^{(l)}, x_i\}_{i=1}^n$ ,  $l = 1, \dots, m$ , constructed from  $m$  independent random samples of size  $n$  from a uniform distribution. The average-jittering estimator is formally specified as

$$\hat{\gamma}_m^A(\alpha) = \frac{1}{m} \sum_{l=1}^m \hat{\gamma}^{(l)}(\alpha)$$

where  $\hat{\gamma}^{(l)}(\alpha)$  is the QR estimator based on  $\{y_i + u_i^{(l)}, x_i\}$ .

Letting subscript  $j$  denote the  $j$ th element of  $x$ , the coefficient  $\gamma(\alpha)$  can be expressed as follows

$$\frac{\partial Q_z(\alpha | x_i)}{\partial x_{ij}} = \exp[x_i' \gamma(\alpha)] \gamma_j(\alpha)$$

hence the usual interpretation of the coefficients in a standard non-linear model applies. In other words, holding everything else constant, the change needed in  $x_{ij}$  to induce a change in the conditional quantile of  $y_i$  of one unit is inversely proportional to  $\gamma_j(\alpha)$ .

### III.2 Raw Family Size Effect

In this section we present the result of QR on counts. These parsimonious specifications include as controls only the origin-family size and a constant. Figure 3 shows the estimated coefficients across the entire distribution of completed fertility, where the estimates with and without zero



counts are presented separately. The grey dotted line represents the effect of origin-family size on all women, with and without children. Notice the effects are initially declining and then increasing. The larger effect is found at the beginning, as well as at the end of the conditional distribution. Now refer to the black solid line, which is the estimated coefficient of family size in the sample containing only women with children. This shows that, with the zeroes removed from the sample, the initial peak for the full sample disappears, and the impact of origin-family size now increases steadily from the 40<sup>th</sup> percentile and reaches its peak at around the 90<sup>th</sup> percentile. This highlights two interesting findings. First, origin-family size is a very important factor in explaining completed fertility for both women with and without children. Secondly, for all women with children, we found that not only is the magnitude of family size on fertility not constant across the entire distribution, but it increases as we move towards the upper tail of the conditional distribution. This means that the effect of origin family size is larger for women with more children in the destination family. All these variations would be hidden if analysis were undertaken using a mean regression.

We summarize the result of Figure 3 in Table 4. We report the estimated family size coefficient at the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentile, along with the estimated marginal effects. Panel A gives the estimates for all 2103 women, and Panel B reports the estimates when only the 1903 women with positive counts are included. For each QR of count data we select 2000 jittered samples.<sup>22</sup> We also report the results from a Poisson regression for comparison, even though it was rejected in the Chi-squared test as discussed in the previous section.<sup>23</sup> Notice that all the estimates are positive and statistically significant regardless which methodology was used. The estimated marginal effects, taken from Panel B (all women with children) as an

---

<sup>22</sup> We started with  $m=2000$  jitters. This procedure is followed iteratively by adding 100 extra number of jittered sample to the model. We repeated this procedure until  $m=2500$  jitters and no significant changes in parameters and standard errors were detected. To end this we report only the result of jittered samples of 2000.

<sup>23</sup> We also tried modelling fertility using Ordinary Least Squares (OLS) estimation. The magnitude and significance of variables in OLS estimation is very similar to that in the Poisson model. Therefore, in the interests of space, we do not report the OLS estimates here.

example, show that a unit increase in the origin-family size from the sample mean leads to an increase of 0.089 units in the conditional mean of fertility by Poisson regression. The marginal effects of origin family size by QR, however, indicate that there is a variation across the different regions of the fertility distribution, ranges from 0.022 to 0.153.<sup>24</sup>

### **III.3. Estimates of Probability of Having Had any Children**

We initially present, in Table 5, estimates from a simple probit model of the probability of having had any children. The dependent value takes the value zero when the respondent has no children, and the value unity when she has one or more children. Origin-family size is negatively correlated with the probability of having one child or more, but this is not statistically significant at conventional levels. The only significant variables are the age cohort dummies and the age at first marriage. Older women are more likely to have had a child than the base group of women age 45 to 50, while women who were older at first marriage are less likely to have reproduced. These findings are consistent with the hypotheses that peer effects matter or that contraceptive technology also matters. The period of fecundity, as proxied by age at first marriage, also has a significant effect.

Our data comprise a low number of childless women (9%), and our estimates above show that the most significant determinants of having children are age cohort and age at first marriage. Neither family size nor the birth order index were statistically significant. As we noted earlier, childlessness is likely to be determined not just by choice but also by the constraints of infertility, especially for our sample of older women (who would mostly not have had access to IVF procedures currently available to younger women). Since we do not wish to include women affected by such constraints in our estimation, and since we have no information about who they

---

<sup>24</sup> To calculate marginal effects, all dummies are set to zero and the continuous variables are set to their means. In other words, the marginal effect can be interpreted the change in the conditional quantile of interest that is induced by a change in the explanatory variable (either dummy or mean), holding all other variables constant.

are, we decided to drop all childless women from our estimation. However the results of QR estimation, using the larger sample of all women, are available from the authors upon request.

#### **III.4. Preferred Model: Estimates with Controls for Family Background Only**

In Table 6a we expand the base specification by QR of count models to include the birth order index, demographics, and family background attributes. We report estimates only for the 1903 women with positive counts for completed fertility.<sup>25</sup> Notice that the origin-family size variable is statistically significant in the Poisson counts model and also for most the quantiles of the conditional fertility distribution.<sup>26</sup> The impact of this variable is found to be increasing across the distribution, from 0.011 at the 10<sup>th</sup> percentile up to 0.043 at the 90<sup>th</sup> percentile. This result indicates that the shape of the conditional distribution, and not just its location, is affected by origin-family size. Notice also that the impact of origin-family size is more precisely estimated from the middle to the upper tail of the fertility distribution than at the lower tail. The marginal estimates of this variable are reported in Table 6b, and they range from about 0.017 in the bottom two quantiles up to 0.131 at the 90<sup>th</sup> quantile. In other words, origin-family size plays a more relevant role in the change of larger counts, say from 7 to 8, than in the change of lower counts, say from 1 to 2.

The sign of the coefficient to the birth order index is negative, as expected. Children of later birth order are less likely to have larger families. This suggests that there are indeed differences in the transmission of fertility norms for children of lower birth order. However, this effect is not statistically significant. Notice that another important determinant in explaining women's fertility outcome is women's age at first marriage. The negative sign implies that women who

---

<sup>25</sup> We began with the larger sample of all women. A comparison of the two models indicated that there were some differences in terms of the magnitude of the coefficients, but only at the lower end of the conditional fertility distribution. The results at the upper tail are remarkably similar.

<sup>26</sup> The only exception is the 10<sup>th</sup> percentile.

were older at first marriage are more likely to have fewer children, and this effect is fairly constant across the conditional completed fertility distribution.<sup>27</sup> Finally, women who grew up in families in which both biological parents were present are less likely to have large families, *ceteris paribus*. This effect is hill-shaped across the conditional completed fertility distribution, and is statistically significant at both the lower and upper end of the conditional fertility distribution. At the 90<sup>th</sup> percentile, the marginal effect, reported in Table 6.1, is -0.511, while at the 50<sup>th</sup> percentile it is only -0.127, and at the 10<sup>th</sup> percentile it is -0.193. Thus the shape of the conditional distribution, and not just its location, is affected by origin-family-normal.

We included birth-cohort and NESB to pick up peer effects. The *ceteris paribus* effect of NESB is more pronounced in the upper tail of the distribution, at the 90<sup>th</sup> percentile, and it is very precisely estimated here. At other parts of the distribution it is insignificantly different from zero. This can be seen as evidence that coming from a NESB is very important in explaining fertility outcome of a woman with more children, but is much less important in determining the fertility decision of women with fewer children in the destination family. The effects of the age-cohort dummy variables also vary across the distribution, and are precisely estimated only at some quantiles.<sup>28</sup>

Empirical results presented in this section show that additional insights were revealed by using a QR of count data which allows for the effect of regressors to differ across different regions of the conditional distribution. The effect origin-family size is very important in determining fertility outcome, and the effect is larger for women with more children in the

---

<sup>27</sup> It is possible that “age at first marriage” might be endogenous, since women whose preferences favour childbearing may marry earlier. To see if our result is affected by endogeneity, we ran our preferred model excluding age at first marriage. We found our result was robust; the significance and magnitude of origin-family size did not alter after the exclusion.

<sup>28</sup> Some might argue that age cohorts are not sufficient to capture cohort effects. We therefore experimented with running the QR regression stratified by different age cohorts. The result suggest that, while the effect of origin family size is slightly more pronounced for older women than younger women in our sample, this variable is statistically significant for each sub-sample.

destination family. These variations would have been hidden by using the fully parametric models.

### **III.5. The Estimates with Controls for Family Background and own Education**

More highly educated women are typically more mobile and may therefore have been more exposed to a variety of peer influences. They are also likely to have greater earning power, and hence their opportunity cost of having children will have been higher. In this subsection we address two main issues with regard to education. First, does any estimated correlation between destination-family size and origin-family size work through omitted human capital effects rather than through cultural transmission? Secondly, does the cultural transmission of fertility norms affect more highly educated women less than lower educated women?

We noted in the Introduction that a challenge for studies attempting to measure the impact of culture on fertility is that there may be some unobserved characteristic that is correlated with both the cultural proxy (origin-family size) and with the independent variable (completed fertility). For example, highly educated parents in the origin family are less likely to have had large families, and highly educated parents are also more likely to have higher educated children. Hence any estimated correlation between destination family size and origin family size might work through omitted human capital effects rather than through cultural transmission.

In this sub-section we report our estimates of completed fertility in which we also control for the highest educational level of the woman in the destination family.<sup>29</sup> In particular we wish to

---

<sup>29</sup> Since human capital in the destination family may be partly endogenous, we did not include any human capital variables in our initial specifications reported in Table 6a. Young women who anticipate having a large family in the future may be less likely to invest in human capital as young adults, since they will be less likely to gain the returns through continuous labour market participation. However, we did include controls in all our specifications for origin-family background variables, such as parental education and whether or not the household had many books. These predetermined variables control to a considerable extent for children's subsequent educational attainment, as shown in Booth and Kee (2005).

see if the coefficient of origin-family size alters in this augmented specification. The results are presented in Panel A of Table 7. They show that, while more educated women have lower levels of completed fertility at the upper part of the distribution, nonetheless the effect of origin-family size on the conditional fertility distribution is remarkably similar to the estimates in Table 6. This suggests that our estimated coefficients to our variables of interest – origin-family size – are not suffering from omitted variable bias. Once more we find that the impact of this variable is found to be increasing across the distribution, from 0.02 at the 10<sup>th</sup> percentile up to 0.043 at the 90<sup>th</sup> percentile. Notice also that the impact of origin-family size is now more precisely estimated at the bottom tail of the fertility distribution than it was found to be in the first Table 7.

The *ceteris paribus* negative effect of higher levels of education is found only in the middle and upper tail of the distribution, at the 90<sup>th</sup> percentile. It is very precisely estimated at the 10<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles, as the reported t-statistics make clear. At the 75<sup>th</sup> and 90<sup>th</sup> percentiles the marginal effects are -0.282 and -0.367 respectively. In other words, this implies that having higher education decreases the 90<sup>th</sup> conditional quantile of jittered data by 0.367 units. Thus the shape of the conditional fertility distribution – as well as the location - is affected by higher education.

The second issue with regard to education is this: does the cultural transmission of fertility norms affect higher educated women less than lower educated women? To answer this question, we stratify our original sample into two subsamples, one comprising women with higher levels of education (consisting of other higher qualifications, degree or above), while the other sub-sample comprises the rest. The results are presented in Panel B and C of Table 7. The quantile regression results show that, while the magnitude of the effect of origin-family size is fairly similar across the two subsamples, the coefficient are more precisely determined in the larger sample of women without higher education. This is perhaps not surprising, since there are

1260 women without and only 643 with higher education. However, there are some small differences across subsamples. For example, for highly educated women, the marginal effect of origin-family size at the 10<sup>th</sup> percentile is now 0.052, while it is only 0.031 for the pooled sample. And for the less highly educated subsample, the effect at the 10<sup>th</sup> percentile has dropped to 0.016. We do not wish to push interpretation of these differences too far, however, as the sample size for more highly educated women is relatively small.

In summary, the results reported in Table 6 show that the estimated correlation between destination-family size and origin-family size does not seem to work through omitted human capital effects. The effect of origin-family size on the conditional fertility distribution is fairly similar across specifications with and without the human capital controls. It would therefore seem to be the case that there is cultural transmission of origin-family fertility norms to the next generation.

### **III.6. What Role Does Partner's Origin-Family Play?**

Women typically do not make fertility decisions in isolation. They have partners whose preferences also need to be accommodated.<sup>30</sup> We take this into consideration by linking women with their husbands using the unique household relationship identifier provided in the BHPS. Since we need to use the origin-family characteristics of the woman's partner as well as her own at the time when the fertility decision were made, we have to drop multiply-married and currently single women from this part of our analysis. This is because we do not observe *earlier* partners' family background data. Thus our sub-sample here comprises 1097 continuously

---

<sup>30</sup> A potential endogeneity problem might arise if there is assortative matching by preferences for family size (women who want a large family might partner with men who also want a large family). However, the within-partnership origin family size correlation is only 0.1965. This suggests that this endogeneity problem is unlikely to be severe in our sample.

married women and their husbands whose origin-family information is observed at wave 13.<sup>31</sup> For these continuously partnered households we can include as controls the origin-family size for each spouse. We proxy the weight of the preferences by controlling for the age differences between partners in a as well as for educational differences. We wish to see if the fertility customs in the origin-family of the older partner, or the more educated partner, carries more weight in affecting completed fertility in the destination family.<sup>32</sup>

Table 8 presents the summary statistics of the male partner for our 1,472 continuously partnered couples. The mean origin-family size of men is 3.618 and the mean age of first marriage is 26.02. The table also shows the age differences between men and women in this smaller dataset, which could affect relative bargaining power within the household. There are about 11% of the couples in our sample at the same age, and around 51% of men older than their partner by 1-4 years old. In our sample men are better educated compared to women, about 42.6% of men have at least other higher qualification, degree or above, and only about 27% of men have no qualification or undefined qualification.

The results controlling for partner's characteristics are reported in Panel A of Table 9. The impact of the wife's origin-family size is now fairly flat in the bottom tail and the middle of the distribution, at around 0.02, and increasing to about 0.034 at the 75<sup>th</sup> and 90<sup>th</sup> percentiles. Now consider the impact of the husband's origin-family size. This is non-linear, being fairly flat at the bottom, increasing to 0.043 at the 75<sup>th</sup> percentile and falling to 0.028 at the 90<sup>th</sup> percentile. This result again indicates that the shape of the conditional distribution, and not just its location, is affected by origin-family size. Notice also that the impact of origin-family size is typically more precisely estimated from the middle to the upper tail of the fertility distribution than at the

---

<sup>31</sup> We also excluded zero counts in this section for the same reasons outlined in the previous section.

<sup>32</sup> We initially estimated the model including age-differences dummies and educational difference controls. We did not find any convincing evidence that the educational differences matter (none of the controls were statistically significant) and hence we do not report these estimates in the paper.



lower tail. The marginal estimates of this variable are reported in Table 9.1, and they range from about 0.026 in the bottom two quantiles up to 0.84 at the 90<sup>th</sup> quantile. This highlights the fact that not only does the wife's origin-family size matter, but that the husband's origin-family size also plays a very important role in determining the fertility outcome. Moreover, the effect is non-linear.

Two other important factors determining fertility are age at marriage and NESB. Here age at first marriage also has the expected negative sign, as previously found in our preferred model. It is highly statistically significant throughout the entire conditional fertility distribution, with marginal effects ranging from -0.038 to -0.113. Wife is NESB has a positive effect on fertility whilst husband is NESB affects fertility negatively from the 10<sup>th</sup> to 75<sup>th</sup> quantile. However, it is not statistically significant at the lower tail at all levels. This result is not surprising as only about 0.007 of men in our sample are NESB.

Panel B of Table 9 extends this model by incorporating partner's age difference dummies. We found the effects of origin-family size of both husband and wife are remarkably similar to the estimates to the results in Panel A. This finding once again justifies the robustness of our model. In terms of age difference dummies, most of the dummy variables are not significant at any statistical levels. In short by using age differences as proxy, we did not find evidence that weight of preferences in affecting completed fertility in the destination family.

#### **IV: THE IMPORTANCE OF CHILDREN TO THE THIRD-GENERATION**

We also explore attitudes to the importance of children of the smaller sub-sample of the third generation in the BHPS – the children aged 16-30 whose parent's fertility history has been observed. Since we can match these children with their parents, we thus know the completed fertility across two families as well as the children's fertility preferences of the third generation.

We link the children with their child-mother by using the household relationship identifier, in a way that is similar to how we relate females to their partners in the previous section.<sup>33</sup> In total there are 1,360 children aged between 10 to 30 years old provided valid response to the survey. Table 10 presents summary statistics of the children in the destination family. Moreover in wave 13 the variable *MLFIMPC* of question MV26C asked respondents about what they think of the importance of having children, where ‘1’ equals “not important at all” and ‘10’ equals “very important”. We used children’s response to this question as an approximation to investigate the evidence of intergeneration fertility pattern.<sup>34</sup> The mean of this variable in the sub-sample is 6.13.

Table 11 reports the result of fertility preferences in the third generation by using an ordered probit model. To begin we use only destination family size as the control variable, and we gradually introduced additional controls such as birth order index and family background to the model. Family-size variable is not significant at all statistical level in all models. Interestingly, in contrast to the previous model, we found that birth-order index is statistically significant in explaining fertility preference in the third generation. The positive coefficient of around 0.17 implies that third generation with higher birth order ranking regard children as more important than children with lower birth order. Parents’ education variables have a negative sign, but are not statistically significant. Finally we also controlled for the presence of books as a proxy of family-specific attitudes to education. We found that, compared to the base group of children with fewer books during childhood, the presence of books affects the attitudes of children in the third generation positively.

---

<sup>33</sup> Note that only those children who live with their mothers and those who responded to the interview were included in this sub-sample.

<sup>34</sup> There is some controversy – and indeed scepticism – in the literature about the quality and meaning of fertility preferences data. Miranda (2005) provides an excellent literature review on this issue. We recognize the limitations of using planned fertility preferences as the dependent variable, but in the absence of any better measure for our 3<sup>rd</sup>-generation, we chose in this section to use it as a proxy of intergenerational fertility patterns.

## V. CONCLUSION

Recent studies by economists exploring the nexus between culture and fertility have modelled culture in the country of origin rather than the family of origin. Our paper extended this avenue of research by investigating how *family-specific* cultural transmission can affect fertility rates. We also allowed for peer-group influences through the inclusion of controls for age cohorts and for country of birth. In line with the previous findings of Fernandez and Fogli (2005), we find that our ‘cultural proxies’ have positive and significant effect in terms of fertility outcomes and thus should not be ignored. Using count data and quantile regression techniques, we find that a woman’s origin-family size is positively correlated with her own completed fertility in the destination family. For our smaller sub-sample of continuously partnered men and women, for whom we are able to estimate the effect of both partners’ origin-family sizes, we find that both significantly affect destination-family fertility.<sup>35</sup>

Debates about declining fertility rates in developed countries have focused on averages. But our results indicate that the variance of completed fertility across families can also have an effect on future fertility within any one country. For example, men and women whose origin-families were large are significantly more likely to have larger families themselves. This pattern is likely to feed through to the subsequent generation. If two countries have identical mean fertility rates but different variances, it is obvious that the one with the higher variance will converge slightly more slowly to lower fertility rates because of the inter-generational transmission of fertility norms.

---

<sup>35</sup> A corollary is that that origin-family size might make a potential instrument for fertility in estimation of models in which fertility is endogenous.

## References

- Anderton, D.L., NO Tsuya, LE Bean, GP Mineau (1987), "Intergenerational Transmission of Relative Fertility and Life Course Patterns," *Demography*, 24(4), 467-480.
- Axinn, WG, ME Clarkberg and A Thornton (1994), "Family Influences on Family Size Preferences," *Demography*, 31(1), 65-79.
- Becker, Gary S., (1960). "An Economic Analysis of Fertility," *Demographic and Economic Change in Developed Countries*, Gary S. Becker, (ed.) Princeton, NJ: Princeton University Press.
- Becker, Gary S., and Robert J. Barro, (1988). "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, pp. 103, 1-25.
- Becker, Gary S., and H. Gregg Lewis, (1973). "On the Interaction Between the Quantity and Quality of Children," *Journal of Political Economy*, 81, S279-S288.
- Becker, Gary S., and Nigel Tomes, (1976). "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84(4) Part 2, S143-S162.
- Bisin, A. and Verdier, T. (2001), 'The Economics of Cultural Transmission and the Dynamics of Preferences', *Journal of Economic Theory* 97(2), 298-319.
- Bisin, A., and T. Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics* 115(3), 955-988.
- Black, Sandra E, Paul J Devereux and Kjell G. Salvanes (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education". *Quarterly Journal of Economics*, 120(2), May.
- Blau, F. D. (1992), "The Fertility of Immigrant Women: Evidence from High Fertility Source Countries," in: Borjas, G. J., Freeman, R. (Eds.), *Immigration and the Workforce: Economic Consequences for the United States and Source Areas*, Chicago: University of Chicago Press, 93-133.
- Booth, A. L. and Kee, H. J. (2005), 'Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment', *IZA Discussion Papers* 1713.
- Borjas, G. (1992), "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics* 107(1), 123-150.
- Borjas, G. (1995), "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review* 85(3), 365-390.
- Bowles, S and H Gintis (2002). "The Inheritance of Inequality". *Journal of Economic Perspectives*, Vol. 16(3), pp. 3-30, Summer.

- Cameron, A.C. and Trivedi, P.K. (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests," *Journal of Applied Econometrics* 1(1), 29-53.
- Conley, Dalton, (2000). "Sibling Sex Composition: Effects on Educational Attainment," *Social Science Research*, 24, pp. 441-457.
- Dunne, M. P., Martin, N. G., Statham, D. J., Slutske, W. S., Dinwiddie, S. H., Bucholz, K. K., Madden, P. A. F. and Heath, A. C. (1997), 'Genetic and environmental contributions to variance in age at first sexual intercourse', *Psychological Science* 8(3), 211-216.
- Efron, B. (1992), "Poisson overdispersion estimates based on the method of asymmetric maximum likelihood," *Journal of the American Statistical Association* 87(417), 98-107.
- Ejrnaes, Mette and Portner, C.C. (2004), "Birth Order and the Intrahousehold Allocation of Time and Education," *Review of Economics and Statistics* LXXXVI(4), 1008-1019.
- Elster, J. (1989), "Social Norms and Economic Theory," *Journal of Economic Perspectives* 3(4), 99-117.
- Ermisch, J.F. (2003), *An Economic Analysis of Family*, Princeton: University Press.
- Fernández, R. and A. Fogli (2005) "Culture: An Empirical Investigation of Beliefs, Work and Fertility," Federal Reserve Bank of Minneapolis, Research Dept. Staff Report 361, April.
- Fernández, R., A. Fogli, and C. Olivetti (2004), "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," *Quarterly Journal of Economics*, 119(4), 1249-1299.
- Guinnane, T., C. Moehling, and C. O Grada (2002), "The Fertility of the Irish in America in 1910," mimeo, Yale University.
- Hendershot, G.E. (1969) "Familial Satisfaction, Birth Order, and Fertility Values," *Journal of Marriage and the Family* 31(1), 27-33.
- Hanushek, Eric A., (1992). "The Trade-off between Child Quantity and Quality," *Journal of Political Economy*, Vol. 100, No. 1. (Feb., 1992), pp. 84-117.
- Huber, P. (1981), *Robust Statistics*, New York: Wiley.
- Kohler, H.-P., Rodgers, J. L. and Christensen, K. (1999), 'Is fertility behaviour in our genes? Findings from a Danish twins study', *Population and Development Review* 25(2), 253-288.
- Lee, M. (1992), "Median regression for ordered discrete response," *Journal of Econometrics* 51(1-2), 59-77.
- Lee, R. (2003), "The demographic Transition: Three Centuries of Fundamental Change". *Journal of Economic Perspectives*, 17, 167-190.

- Livi-Bacci, M. (2001), *A Concise History of World Population*, London: Blackwell.
- Machado, J. A. F. and Santos Silva, J. M. C. (2005), 'Quantiles for Counts', *Journal of the American Statistical Association* 100(472), 1226-1237(1212).
- Melkersson, Maria and Rooth, Dan-Olof. (2000), "Modeling female fertility using inflated count data models," *Journal of Population Economics* 13(2), 189-203.
- Miller, W. B., Pasta, D. J., MacMurray, J., Chiu, C., Wu, S. and Comings, D. E. (1999), 'Genetic Influences on Childbearing Motivation: A Theoretical Framework and some Empirical Evidence', *Advances In Population* 3, 53-102.
- Miranda, Alfonso. (2005), "Planned Fertility and Family Background: A Quantile Regression for Counts Analysis," Keele Economics Research Papers KERP 2005/07.
- Miranda, Alfonso, (2003). "Socio-Economic characteristics, completed fertility and the transition from low to high order parities in Mexico," EconWPA Discussion Paper no: 0308001.
- Murphy, Michael and Knudsen, L. B. (2002), "The Intergenerational Transmission of Fertility in Contemporary Denmark: The Effects of Number of Siblings (Full and Half), Birth Order, and Whether Male or Female," *Population Studies* 56(3), 235-248.
- Murphy, Michael and Wang, D. (2001), "Family-Level Continuities in Childbearing in Low-Fertility Societies," *European Journal of Population* 17(1), 74-96.
- Pearson, K. and Lee, A. (1899), 'On the Inheritance of Mankind', *Royal Society of London Philosophical Transactions Ser A*. 192.
- Rodgers, J. L., Kohler, H.-P., Kyvik, K. O. and Christensen, K. (2001), 'Behavior genetic modelling of human fertility: findings from a contemporary Danish twin study', *Demography* 38(1), 29-42.
- Santos Silva, J. M. C. and Covas, Francisco. (2000), "A modified hurdle model for completed fertility," *Journal of Population Economics* 13(2), 173-188
- Stevens, W.L. (1950), "Fiducial limits of the parameter of a discontinuous distribution," *Biometrika* 37(1-2), 117-129.
- Sulloway, F.J. (1996). *Born to Rebel: Birth Order, Family Dynamics, and Creative Lives*, New York: Pantheon, 1996.
- Wang, W. and Famoye, F. (1997), "Modeling Household Fertility Decisions with Generalized Poisson Regression," *Journal of Population Economics* 10(3), 273-283.
- Westoff, C. F. and Potvin, R. H. (1967), *College Women and Fertility Values*, Princeton, NJ: Princeton University Press.

Winkelmann, R. (2003), *Econometric Analysis of Count Data*, Berlin Heidelberg, New York: Springer Verlag, 4th Ed.

Winkelmann, R. (1995), "Duration Dependence and Dispersion in Count-Data Models," *Journal of Business & Economic Statistics* 13(4), 467-474.

Winkelmann, R. and Zimmermann, K.F. (1994), "Count Data Models for Demographic Data," *Mathematical Population Studies* 4(3), 205-221.

## Appendix A: Data Appendix

### 1. The Dependent Variable

We provide a summary of the process to construct the dependent variable, *the number of biological children that respondent has ever had*. In the BHPS, fertility related questions were first asked in the second wave, when each respondent was asked (question BL42): “Do you have or have you ever had/fathered any children?”<sup>36</sup> If the answer to this question was “Yes”, respondents were immediately asked the subsequent question (BL43): “How many children have you had/fathered in all?” We derived the dependent variable based on valid responses to the second question.

In each of the subsequent waves, this question is only asked if the respondent is either a new entrant to the survey, or had not been asked the same question before. As a result of this particular data structure, even though we are only interested in the family background data from wave 13, it was necessary to retrieve fertility information from the earliest possible wave. The variables of interest are  $xLPRNT$  and  $xLNPNT$ , both found in wave 2, and also from wave 8 to 13. Each respondent was asked the two questions given in the preceding paragraph. ( $xLPRNT$  of question BL42): “Do you have or have you ever had/fathered any children?”<sup>37</sup> If the answer to this question is “Yes”, respondents were immediately asked the subsequent question ( $xLNPNT$  of BL43): “How many children have you had/fathered in all?” Conditional on the response is yes to the first question, we derived the dependent variable based on the valid response in the second question.

However, respondents were only asked this question once. This implies that, for those respondents who were still fertile, any new births will not be captured in the variable  $xLNPNT$ . We therefore incorporated the new births that were reported in the household file  $wINDALL$ . The variable  $xNEWHY$  in the household file reveals the existence of new birth in the household. We thus relate this information to their mother using several key linking variables provided in the household relationship grid. This procedure was repeated in every subsequent wave after wave 2 for each of the existing respondents, to update the fertility variables.

---

<sup>36</sup> This includes biological children and stillbirths only. Exclusions are miscarriages and adopted, fostered or step children. For men, includes children where respondent and the mother never lived together.

<sup>37</sup> This includes biological children and stillbirths only. Exclusions are miscarriages and adopted, fostered or step children. For men, includes children where respondent and the mother never lived together.



## 2. The Birth Order Index

In the raw data, the simple correlation coefficient between family size and birth order is 0.6643. By construction, our index effectively purges family size from birth order and consequently the simple correlation coefficient between family size and the birth order *index* is just 0.0372.

Suppose  $N$  is total number of siblings in the respondent's origin-family including the respondent,  $\phi$  is the absolute birth order of the respondent and  $A$  denotes average birth order in each origin-family. Thus the absolute birth order variable  $\phi$  takes the value 1 for the first born, 2 for the 2<sup>nd</sup> born, and so on, up to a top value of 10 for the 10<sup>th</sup> born and above. "Only" children are assigned the same birth order as first born children. Average birth order  $A$  is calculated as  $(N+1)/2$  and is clearly increasing in origin family size and bounded between 1 and 5.5.<sup>38</sup>

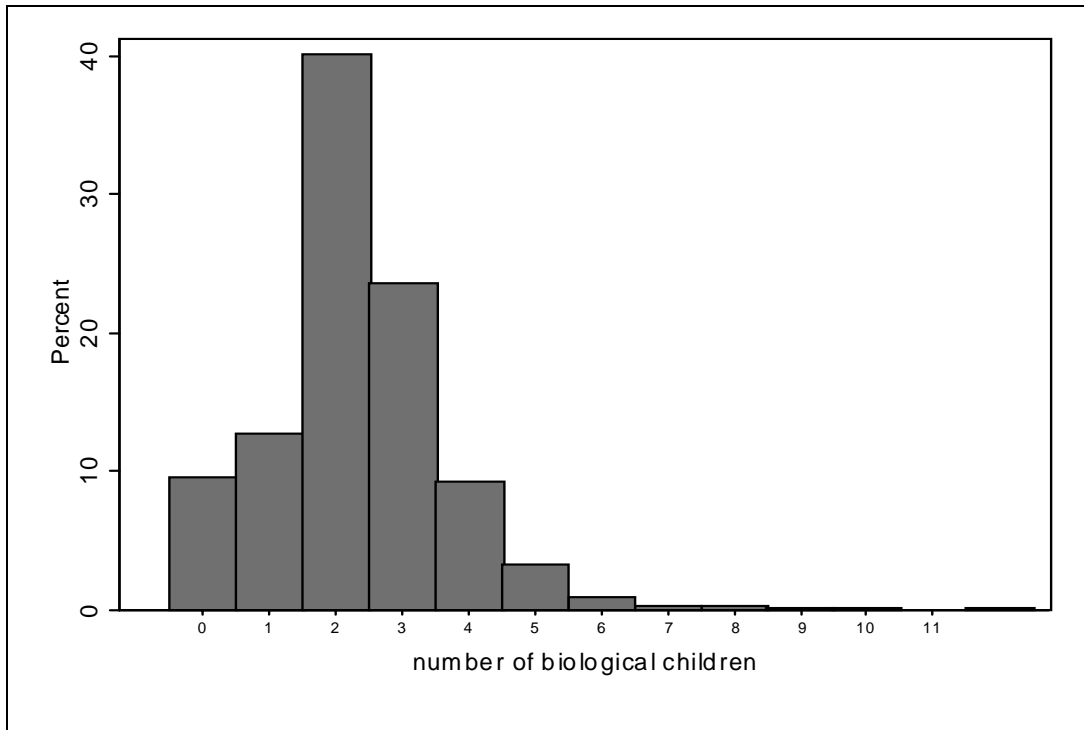
Now let  $B$  denote the birth order index, where  $B = \phi/A$ ; that is,  $B$  is the ratio of the respondent's birth order to the average birth order of her origin family and for our data  $B \in (0.18, 1.82)$ .<sup>39</sup> Importantly, notice that, by construction, the within-family mean of  $B=1$  is the same across all origin-family types. Thus  $B=1$  represents both the within-family and across family mean. Deflating birth order  $\phi$  by average birth order within the origin-family  $A$  ensures that our constructed birth order index  $B$  is independent of origin-family size. Note that the variance is also constant across all origin-family types, as shown in Booth and Kee (2006).

---

<sup>38</sup> For a one-child family, average birth order  $A = 1$ , for a 2-child family,  $A = 1.5$ , for a 3-child family  $A = (3+1)/2=2$ , and so on, up to a total value for the 10-child family of  $A = (10+1)/2=5.5$ .

<sup>39</sup> To illustrate, consider four family types: 1-child, 2-child, 3-child and 10-child. For the only child from a one-child family,  $B_{11}=1$ , where the first subscript denotes birth order and the second family size. Now consider the first born child from a 2-child family. Her index is  $B_{12}=1/1.5=0.666$ . For the 2<sup>nd</sup> born child,  $B_{22}=2/1.5=1.333$ . Next, take a 3-child family. The first born has  $B_{13}=0.5$ , the 2<sup>nd</sup> born has  $B_{23}=1$ , while the 3<sup>rd</sup> born has  $B_{33}=3/2=1.5$ . Finally, consider a 10-child family. Here the first born has  $B_{1,10} = 1/(5.5)=0.182$ , the 2<sup>nd</sup> born has  $B_{2,10}=2/(5.5)=0.364$ , the 3<sup>rd</sup> born has  $B_{3,10}=3/(5.5)=0.545$ , the 9<sup>th</sup> born has  $B_{9,10}=9/(5.5)=1.636$ , while the 10<sup>th</sup> born has  $B_{10,10}=10/(5.5)=1.818$ .

**Figure 1: Empirical Distribution of Number of Biological Children**



**Table 1: Observed and Estimated Poisson distribution  
(with sample mean of 2.27 children)**

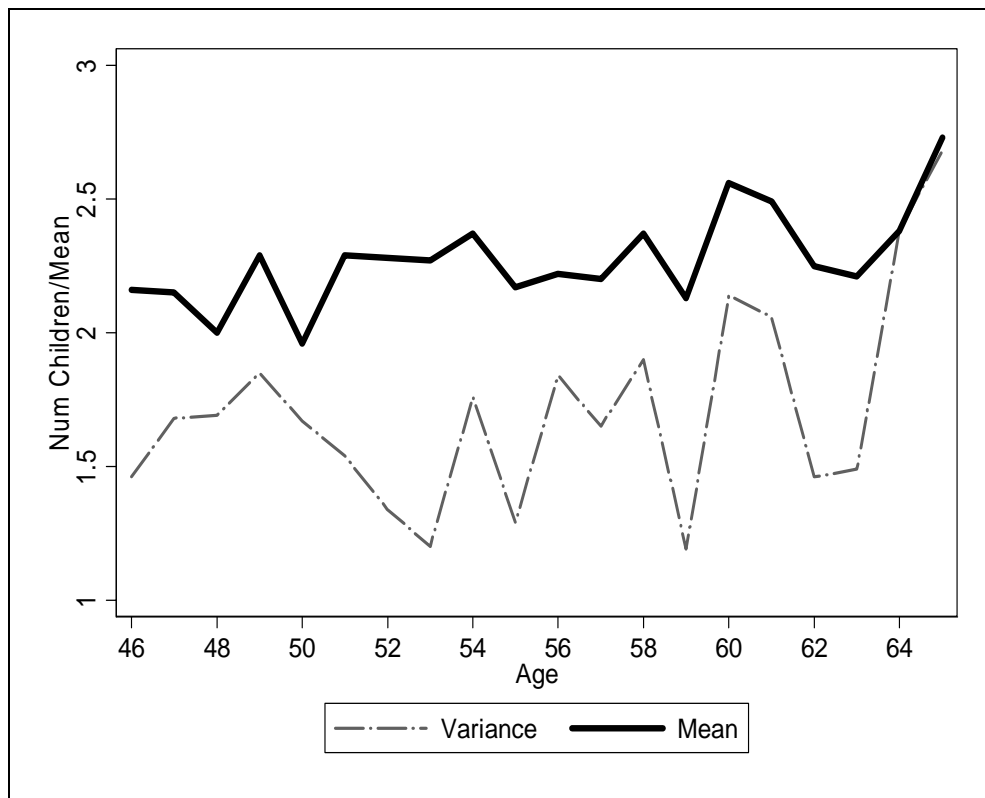
	No.obs	% Share	Poisson %
0	200	9.51%	10.32%
1	269	12.79%	23.43%
2	846	40.23%	26.61%
3	496	23.59%	20.15%
4	193	9.18%	11.44%
5	67	3.19%	5.20%
6	18	0.86%	1.97%
7	5	0.24%	0.64%
8	5	0.24%	0.18%
9+	4	0.19%	0.09%
N	2103		
$\chi^2$			322.06

Note: 1. For raw data: Mean=2.27, Variance=1.73.  
2.  $\chi^2_{10}=18.31$  at the 5% level.

**Table 2: Completed Fertility by Origin-Family Size (Women aged 45-65 in 2003, %)**

Family Size	Completed Fertility: Number of Biological Children											Total %/(n)	
	0	1	2	3	4	5	6	7	8	9	>=10		
1	15.2	13.2	44	18.8	6.8	1.6	0.4	0	0	0	0	11.89	250
2	10.76	15.54	42.83	22.71	7.17	0.6	0.2	0	0	0.2	0	23.87	502
3	8.06	13.13	41.01	24.88	8.99	2.76	0.46	0.46	0	0	0.23	20.64	434
4	6.62	13.25	36.28	27.76	10.09	4.73	0.63	0	0.63	0	0	15.07	317
5	7.49	10.7	43.85	25.67	9.09	1.6	1.07	0	0.53	0	0	8.89	187
6	13.33	9.63	39.26	23.7	7.41	2.22	2.22	0.74	0.74	0.74	0	6.42	135
7	5.68	4.55	34.09	28.41	17.05	6.82	2.27	1.14	0	0	0	4.18	88
8	9.23	18.46	29.23	16.92	15.38	6.15	1.54	1.54	1.54	0	0	3.09	65
9	5.66	15.09	39.62	18.87	7.55	9.43	3.77	0	0	0	0	2.52	53
>=10	8.33	2.78	31.94	18.06	18.06	16.67	2.78	0	0	0	1.39	3.42	72
<b>Total % (n)</b>	9.51 200	12.79 269	40.23 846	23.59 496	9.18 193	3.19 67	0.86 18	0.24 5	0.24 5	0.10 2	0.10 2	100	2103

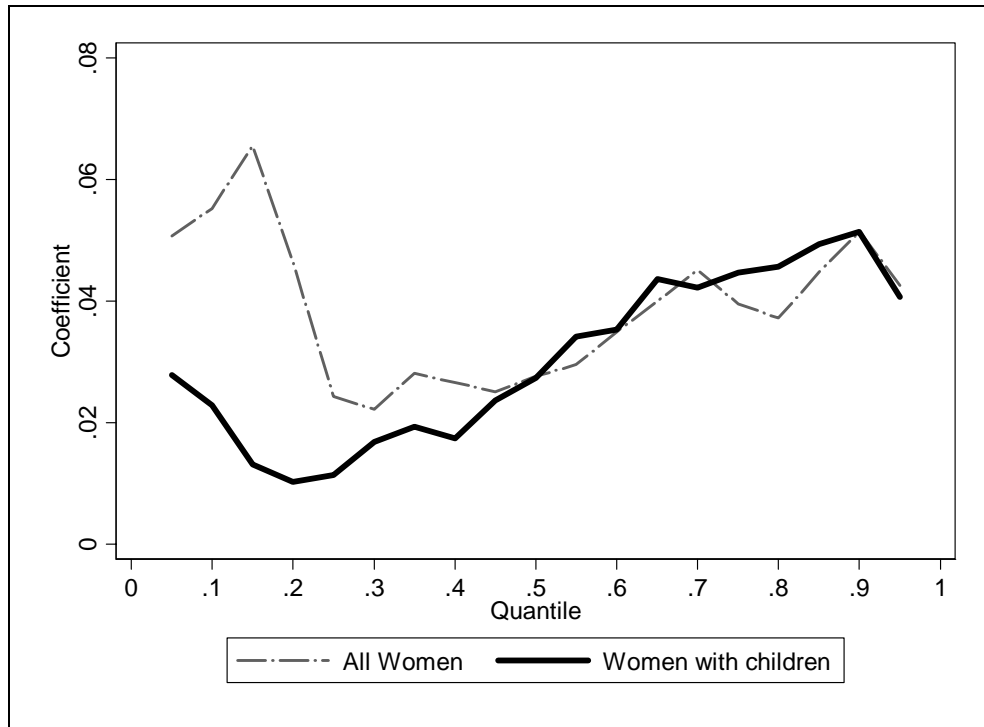
**Figure 2: Mean and Variance of Completed Fertility, Women aged 45-65**



**Table 3: Summary Statistics of Child-mother in the Origin Family**

<b>Name</b>	<b>Mean</b>	<b>Description (N=2,103)</b>
<b><u>Family Composition</u></b>		
Family size	3.757	Number of children in respondent's own family, top-coded at 10.
logB2	-0.127	Log B-Order index.
<b><u>Demographics</u></b>		
NESB	0.008	Dummy=1 if respondent comes from a Non-English-Speaking-background. (ie English was not 1 <sup>st</sup> language)
Other Eng-sp countries	0.016	Dummy=1 if respondent was born in other English speaking country: Australia, New Zealand, Canada, Ireland, South Africa and the United States.
Age 46-50	0.254	Dummy=1 if respondent aged between 46-50 years old.
Age 51-55	0.266	Dummy=1 if respondent aged between 51-55 years old.
Age 56-60	0.277	Dummy=1 if respondent aged between 56-60 years old.
Age 61-65	0.203	Dummy=1 if respondent aged between 61-65 years old.
Age at first marriage	23.96	Age at first marriage.
married	0.742	Dummy=1 if respondent is currently legally married.
cohab	0.041	Dummy=1 if respondent is currently cohabiting.
marr_oth	0.217	Dummy=1 if respondent is widowed, divorced, separated or never married.
Edu_low	0.325	Dummy=1 if respondent has no defined qualification currently studying or educational information is missing.
Edu_norm	0.329	Dummy=1 if respondent has some schooling qualification, O or A level.
Edu_high	0.346	Dummy=1 if respondent has other higher qualification, degree or above.
<b><u>Family Attributes</u></b>		
Lots of books	0.315	Dummy=1 if respondent had lots of books during childhood.
Quite a few books	0.346	Dummy=1 if respondent had quite a few books during childhood.
Less book	0.329	Dummy=1 if respondent had not many books during childhood.
Father education	0.323	Dummy=1 if respondents' father has some qualification, further education, degree or further qualification.
Mother education	0.236	Dummy=1 if respondents' mother has some qualification, further education, degree or further qualification.
Dad age when Born	27.56	Respondents' father age when respondent was born.
Mum age when born	26.25	Respondents' mother age when respondent was born.
Mum worked	0.378	Respondent's mother worked when respondent was age 14.
Family normal	0.864	Dummy=1 if respondent lived with both biological parents until age 16.
kidinner	0.120	Dummy=1 if respondent lived in inner city during childhood.
kidsuburb	0.219	Dummy=1 if respondent lived in suburban area during childhood.
Kidtown	0.240	Dummy=1 if respondent lived in town during childhood.
Kidvillage	0.194	Dummy=1 if respondent lived in village during childhood.
Kidrural	0.190	Dummy=1 if respondent lived in rural during childhood.
Kidmove	0.037	Dummy=1 if respondent moved around during childhood.

**Figure 3: Raw Family Size Effect for Women With and Without Children**



**Table 4: Parsimonious Specification, Family Size Effect**  
(Number of biological children as Dependent Variable)

	Poisson	Quantile Regression Model				
		10th	25th	50th	75th	90 <sup>th</sup>
<b><u>A. Women with and without children</u></b>						
Family size	0.042 (6.95)***	0.055 (2.11)**	0.024 (3.49)***	0.028 (4.56)***	0.04 (5.89)***	0.051 (9.92)***
Constant	0.659 (23.52)***	-0.284 (-2.31)**	0.508 (13.59)***	0.674 (30.82)***	0.871 (30.54)***	1.036 (35.03)***
Marginal Effects	0.095	0.043	0.041	0.055	0.096	0.149
Observations	2103	2103	2103	2103	2103	2103
<b><u>B. Women with children</u></b>						
Family size	0.035 (5.87)***	0.023 (3.27)***	0.011 (3.06)***	0.027 (4.58)***	0.045 (7.01)***	0.051 (10.40)***
Constant	0.782 (27.78)	0.428 (10.86)***	0.647 (40.00)***	0.735 (36.01)***	0.878 (33.04)***	1.067 (39.54)***
Marginal Effects	0.089	0.036	0.022	0.058	0.110	0.153
Observations	1903	1903	1903	1903	1903	1903

Note: 1 Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

**Table 5: Specification 1- Probability of Having Children**  
(Whether or not Respondent has had children as Dependent Variable)

	<b>Probit</b>
<b><u>Family Composition</u></b>	
Family size	0.032 (1.53)
Log B index	-0.061 (0.57)
<b><u>Demographics</u></b>	
NESB	0.171 (0.32)
Other Eng-sp countries	-0.478 (1.86)*
Age 51-55	0.404 (3.43)***
Age 56-60	0.181 (1.66)
Age 61-65	0.268 (2.12)**
Age at first marriage	-0.065 (8.42)***
<b><u>Family Attributes</u></b>	
Lots of books	0.017 (0.15)
Quite a few books	-0.1 (0.96)
Family normal	-0.027 (0.2)
Mum worked	0.031 (0.34)
Father education	-0.127 (1.23)
Mother education	0.005 (0.05)
Father age when R was born	-0.007 (0.85)
Mother age when R was born	0.005 (0.48)
Constant	2.871 (6.91)***
<hr/>	
Area when young Dummies <sup>3</sup>	YES
Observations	2103
Pseudo-R2	0.1584
Log Likelihood	-556.07
LR-chi2	209.33

Note: 1. Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. 3. Area when young dummies include kidinner, kidsuburb, kidtown, kidvillage and kidrural, with kidmove as the base of controls.

**Table 6a: Preferred Model-QR Spec (Controls for All Family Background Variables)**  
(Number of biological children as Dependent Variable, Zero Count Excluded)

	<b>Poisson</b>	<b>10th</b>	<b>25th</b>	<b>50th</b>	<b>75th</b>	<b>90th</b>
<b><u>Family Composition</u></b>						
Family size	0.029 (4.35)***	0.011 (1.6)	0.01 (2.14)**	0.028 (4.72)***	0.035 (5.20)***	0.043 (6.38)***
Log B index	-0.040 (1.16)	-0.048 (1.26)	-0.033 (1.35)	-0.008 (0.28)	-0.042 (1.24)	-0.027 (0.86)
<b><u>Demographics</u></b>						
NESB	0.324 (2.39)**	-0.25 (1.23)	0.131 (0.75)	0.209 (1.68)*	0.197 (0.74)	0.331 (2.54)**
other Eng-sp countries	0.157 (1.34)	-0.112 (0.64)	0.02 (0.21)	0.187 (0.79)	0.194 (1.6)	0.095 (0.58)
Age 51-55	-0.008 (0.19)	-0.026 (0.58)	0.006 (0.24)	-0.02 (0.74)	0.044 (1.07)	0.01 (0.25)
Age 56-60	0.046 (1.12)	0.039 (0.81)	0.051 (1.94)*	0.061 (2.22)**	0.064 (1.70)*	-0.016 (0.39)
Age 61-65	0.06 (1.36)	0.04 (0.9)	0.039 (1.38)	0.053 (1.76)	0.073 (1.65)*	0.032 (0.65)
Age at first marriage	-0.016 (4.27)***	-0.021 (6.16)***	-0.016 (5.04)***	-0.011 (3.73)***	-0.011 (3.37)***	-0.011 (3.30)***
<b><u>Family Attributes</u></b>						
Lots of books	-0.013 (0.34)	0.024 (0.53)	-0.026 (0.98)	-0.061 (2.32)**	-0.051 (1.43)	0.022 (0.58)
Quite a few book	-0.008 (0.21)	0.052 (1.27)	0.017 (0.78)	-0.038 (1.53)	-0.02 (0.67)	-0.025 (0.71)
Family normal	-0.062 (1.39)	-0.121 (2.76)***	-0.069 (2.63)***	-0.058 (1.83)*	-0.069 (1.53)	-0.141 (2.72)***
Mum worked	-0.021 (0.66)	-0.068 (1.70)*	-0.022 (1.08)	0.013 (0.65)	-0.01 (0.34)	-0.018 (0.52)
Fath education	-0.051 (1.36)	-0.098 (2.11)*	-0.031 (1.39)*	-0.033 (1.36)	-0.056 (1.79)*	-0.035 (1.02)
Moth education	0.018 (0.45)	0.027 (0.5)	0.052 (2.02)**	0.027 (1.04)	0.014 (0.38)	-0.005 (0.16)
Fath age when R was born	-0.001 (0.23)	-0.01 (2.40)**	0.000 (0.09)	0.001 (-0.28)	0.000 (0.15)	-0.002 (1.21)
Moth age when R was born	0.002 (0.59)	0.009 (2.28)**	0.002 (0.84)	0.001 (0.21)	0.003 (0.83)	0.004 (1.34)
Constant	1.134 (7.27)***	1.228 (8.14)***	1.048 (9.66)***	1.084 (9.60)***	1.074 (7.96)***	1.4 (8.38)***
Area when young	YES	YES	YES	YES	YES	YES
Observations	1903	1903	1903	1903	1903	1903

Note: 1. Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. 3. Area when young dummies include kidinner, kidsuburb, kidtown, kidvillage and kidrural, with kidmove as the base of controls.

**Table 6b: Preferred Specification Marginal Effects**

	Poisson	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>
<b><u>Family Composition</u></b>						
Family size	0.072	0.017	0.017	0.056	0.087	0.131
Log B index	-0.100	-0.071	-0.058	-0.016	-0.112	-0.092
<b><u>Demographics</u></b>						
NESB	0.951	-0.341	0.252	0.498	0.598	1.348
other Eng-sp countries	0.422	-0.163	0.037	0.438	0.588	0.345
Age 51-55	-0.020	-0.039	0.011	-0.043	0.123	0.036
Age 56-60	0.116	0.061	0.093	0.133	0.181	-0.055
Age 61-65	0.153	0.063	0.071	0.116	0.206	0.111
Age at first marriage	-0.039	-0.042	-0.036	-0.028	-0.033	-0.044
<b><u>Family Attributes</u></b>						
Lots of books	-0.033	0.036	-0.046	-0.130	-0.140	0.076
Quite a few book	-0.019	0.080	0.031	-0.081	-0.056	-0.085
Family normal	-0.158	-0.193	-0.126	-0.127	-0.194	-0.511
Mum worked	-0.053	-0.104	-0.039	0.029	-0.027	-0.063
Fath education	-0.126	-0.150	-0.056	-0.070	-0.155	-0.120
Moth education	0.046	0.042	0.093	0.058	0.039	-0.018
Fath age when R was born	-0.002	-0.015	-0.000	-0.001	0.001	-0.008
Moth age when R was born	0.005	0.015	0.004	0.001	0.008	0.013



**Table 7: Stratification by Education**  
(Number of biological children as Dependent Variable, Zero Count Excluded)

	10th	25th	50th	75th	90th
<b><u>A. All Mothers</u></b>					
Family size	0.02 (3.11)***	0.013 (2.81)**	0.025 (4.36)***	0.033 (4.58)***	0.043 (6.28)***
Edu_high	0.163 (4.04)***	0.052 (1.89)*	-0.039 (1.41)	-0.100 (2.79)**	-0.104 (2.7)**
Edu_norm	0.079 (1.89)*	0.031 (1.19)	-0.053 (1.92)*	-0.078 (2.30)***	-0.102 (2.81)**
Constant	1.22 (8.68)***	1.049 (9.62)***	1.075 (9.25)***	1.114 (7.76)***	1.544 (10.25)***
Other Controls <sup>3</sup>	YES	YES	YES	YES	YES
Observations	1903	1903	1903	1903	1903
<b><u>B. Mothers with Higher Education</u></b>					
Family size	0.043 (4.26)***	0.014 (1.72)*	0.023 (2.21)**	0.044 (2.74)**	0.034 (2.48)**
Other Controls <sup>3</sup>	YES	YES	YES	YES	YES
Observations	643	643	643	643	643
<b><u>C. Mothers without Higher Education</u></b>					
Family size	0.012 (1.69)*	0.013 (2.26)**	0.025 (3.80)***	0.034 (3.90)***	0.042 (6.02)***
Other Controls <sup>3</sup>	YES	YES	YES	YES	YES
Observations	1260	1260	1260	1260	1260

Note: 1. Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. 3. Other Controls refer to all explanatory variables as specified in the Preferred Specification of Table 3.

**Table 7.1: Stratification by Education –Marginal Effects**

	10th	25th	50th	75th	90th
<b><u>All Mothers</u></b>					
Family size	0.031	0.025	0.054	0.085	0.132
Edu_high	0.271	0.104	-0.092	-0.282	-0.367
Edu_norm	0.132	0.062	-0.122	-0.217	-0.355
<b><u>Mothers with Higher Education</u></b>					
Family size	0.052	0.027	0.049	0.115	0.116
<b><u>Mothers without Higher Education</u></b>					
Family size	0.016	0.021	0.048	0.086	0.042

**Table 8: Summary Statistics of Child-father in the Origin Family**

<b>Name</b>	<b>Mean</b>	<b>Description (N=1,472)</b>
PNESB	0.007	Dummy=1 if respondents' partner comes from a Non-English-Speaking background.
Pt_child2	3.618	Number of children in respondent partner's family, top-coded at 10.
Page_1marr	26.02	Age of first marriage.
Pedu_low	0.270	Dummy=1 if respondent has no defined qualification currently studying or educational information is missing.
Pedu_norm	0.304	Dummy=1 if respondent has some schooling qualification, O or A level.
Pedu_high	0.426	Dummy=1 if respondent has other higher qualification, degree or above.
mf10*	0.032	Dummy=1 if male older than female by at least 10 years old.
mf95*	0.155	Dummy=1 if male older than female by 5-9 years old.
mf41*	0.512	Dummy=1 if male older than female by 1-4 years old.
agesame*	0.114	Dummy=1 if male and female are of same age.
fm41*	0.155	Dummy=1 if female older than male by 1-4 years old.
fm95*	0.025	Dummy=1 if female older than male by 5-9 years old.
fm41*	0.008	Dummy=1 if female older than male by 1-4 years old.

*\* Only referring to women who are currently partnered with her first husband.*

**Table 9: Partner's Family Size Effect**  
(Number of biological children as Dependent Variable, Zero Count Excluded)

	10th	25th	50th	75th	90th
<b>A. Control for Partner's</b>					
<b>Family Size</b>					
Wife's family size	0.02 (2.64)***	0.016 (2.79)***	0.019 (2.86)***	0.034 (4.00)***	0.034 (3.29)***
Wife: Log B index	-0.01 (0.29)	-0.034 (1.13)	-0.055 (1.66)*	-0.057 (1.65)*	-0.028 (0.67)
Wife: NESB	0.258 (1.73)*	0.288 (1.75)*	0.241 (2.27)**	0.305 (2.16)**	0.278 (2.03)**
Wife: Other Eng-speaking	-0.129 (0.48)	-0.131 (0.81)	0.186 (1.11)	0.119 (0.61)	0.088 (0.89)
Wife: Age at marriage	-0.029 (4.96)***	-0.014 (3.75)***	-0.017 (5.13)***	-0.025 (6.22)***	-0.016 (2.25)**
Wife: Age 51-55	-0.009 (0.16)	-0.036 (1.18)	-0.041 (1.18)	0.017 (0.35)	0.023 (0.42)
Wife: Age 56-60	0.065 (1.33)	0.017 (0.57)	0.03 (0.87)	0.037 (0.93)	-0.006 (0.12)
Wife: Age 61-65	0.036 (0.67)	-0.004 (0.13)	0.005 (0.13)	0.022 (0.49)	0.016 (0.25)
Husband is NESB	-0.031 (0.26)	-0.072 (0.55)	-0.031 (0.25)	-0.32 (2.97)***	0.855 (5.61)***
Husband's family size	0.016 (2.01)**	0.012 (2.28)**	0.026 (4.04)***	0.043 (6.60)***	0.028 (2.64)***
Other Controls <sup>3</sup>	YES	YES	YES	YES	YES
Observations	1097	1097	1097	1097	1097
<b>B. Control for Age Differences</b>					
Wife: Family size	0.024 (3.00)***	0.015 (2.53)**	0.02 (2.84)***	0.031 (3.59)***	0.033 (4.25)***
Wife: NESB	0.326 (3.53)***	0.285 (1.15)	0.134 (1.2)	0.308 (2.62)***	0.262 (2.19)**
Wife: Age at marriage	-0.023 (3.43)***	-0.015 (3.06)***	-0.021 (5.35)***	-0.021 (4.94)***	-0.017 (2.49)**
Husband is NESB	-0.111 (0.67)	-0.06 (0.45)	-0.005 (0.04)	-0.334 (3.42)***	0.83 (6.13)***
Husband's family size	0.011 (1.56)	0.014 (2.51)**	0.027 (4.27)***	0.045 (6.91)***	0.028 (3.71)***
M>F by at least 10 yrs	-0.315 (2.63)***	-0.048 (0.39)	0.012 (0.1)	-0.037 (0.25)	0.218 (0.97)
M>F by 5-9 yrs	-0.144 (1.62)	0.042 (0.88)	0.138 (2.62)***	0.08 (1.45)	0.056 (0.89)
M>F by 1-4 yrs	0.004 (0.07)	0.035 (0.98)	0.075 (1.67)*	0.018 (0.38)	0.015 (0.29)
F>M by 1-4 yrs	0.043 (0.7)	0.062 (1.44)	0.092 (1.81)*	0.07 (1.29)	0.049 (0.7)
F>M by 5-9 yrs	-0.383 (3.11)***	-0.523 (3.04)***	0.193 (2.70)***	-0.043 (0.61)	-0.263 (3.25)***
F>M by at least 10 yrs	0.067 (0.27)	0.257 (1.87)*	0.183 (1.22)	0.145 (1.03)	0.368 (3.45)***
Other Controls <sup>3</sup>	YES	YES	YES	YES	YES
Observations	1097	1097	1097	1097	1097

Note: 1. Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. 3. Other Controls include birth order index, born Eng\_sp, Moth worked, Fath education, Moth education, Fath age when R was born, Moth age when R was born, Area when young dummies and age of first marriage.

**Table 9.1: Partner's Family Size – Marginal Effect**

	<b>10th</b>	<b>25th</b>	<b>50<sup>th</sup></b>	<b>75th</b>	<b>90th</b>
<b><u>Partner's Family Size Effect</u></b>					
Family size	0.031	0.030	0.040	0.082	0.098
Log B index	-0.017	-0.065	-0.120	-0.148	-0.088
NESB	0.483	0.658	0.611	0.953	1.026
other Eng-speaking countries	-0.200	-0.242	0.460	0.339	0.295
Age at marriage	-0.088	-0.038	-0.056	-0.113	-0.074
Age 51-55	-0.014	-0.070	-0.092	0.046	0.074
Age 56-60	0.109	0.033	0.068	0.100	-0.020
Age 61-65	0.060	-0.008	0.011	0.058	0.052
Partner is NESB	-0.050	-0.137	-0.068	-0.737	4.312
Partner family size	0.026	0.024	0.055	0.104	0.084
<b><u>Control for Age Differences</u></b>					
Family size	0.036	0.027	0.042	0.075	0.095
NESB	0.635	0.634	0.323	0.960	0.962
Age at marriage	-0.063	-0.038	-0.072	-0.090	-0.081
Partner is NESB	-0.174	-0.112	-0.011	-0.759	4.150
Partner family size	0.017	0.026	0.058	0.109	0.084
M>F by at least 10 yrs	-0.453	-0.090	0.027	-0.097	0.779
M>F by 5-9 yrs	-0.226	0.081	0.325	0.221	0.185
M>F by 1-4 yrs	0.006	0.067	0.168	0.047	0.047
F>M by 1-4 yrs	0.071	0.122	0.213	0.191	0.162
F>M by 5-9 yrs	-0.535	-0.804	0.476	-0.112	-0.754
F>M by at least 10 yrs	0.114	0.563	0.451	0.415	1.426

**Table 10: Summary Statistics for Children in the Destination Family**

Name	Mean	Description (N=1,202)
Cimportance	6.130	Importance of having children, value ranges from 1-10.
Cage	20.212	Children's age.
Ct_child2	2.815	Mean family size in the destination family.
Cdad_deg	0.404	Dummy =1 if child-father has at least a degree.
Cmum_deg	0.348	Dummy =1 if child-mother has at least a degree.
lots_bk	0.442	Dummy=1 if children in the destination have lots of books during childhood.
more_bk	0.426	Dummy=1 if children in the destination have quite a few books during childhood.
less_bk	0.128	Dummy=1 if children in the destination have less books during childhood.

**Table 11: Children in the Destination Family – Ordered probit**  
(Importance of having children as Dependent Variable)

	Raw Family Size	B-index	Parent Education	Books
Family size	0.003 (0.12)	0.007 (0.3)	0.007 (0.3)	0.009 (0.36)
Log B index		0.177 (2.60)***	0.18 (2.64)***	0.179 (2.64)***
Fath education			-0.05 (-0.78)	-0.061 (-0.95)
Moth education			-0.006 (-0.09)	-0.008 (-0.12)
Lots of books				0.154 (1.79)*
Quite a few book <sup>3</sup>				0.202 (2.37)**
Observations	1360	1360	1360	1360

Note: 1. Absolute z statistics in parentheses. 2 \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. 3. Base is less\_bk.