

**Is Having Babies Contagious?  
Estimating Fertility Peer Effects Between Siblings**

Ilyana Kuziemko<sup>†</sup>  
Harvard University

June 2006

<sup>†</sup> [kuziemko@nber.org](mailto:kuziemko@nber.org). I would like to thank David Cutler, Claudia Goldin, Ed Glaeser, Caroline Hoxby, Larry Katz, Steve Levitt and Shanna Rose for valuable comments on drafts, as well as participants at the Harvard Labor/Public Finance workshop for comments on an earlier version of the project.

## **Abstract**

Models of fertility have long relied on peer effects to explain the large variance in birth rates across time and place within the developed world. However, little has been done using micro-level data to establish that fertility peer effects actually exist. I find that the probability of having a child rises substantially in the two years after one's sibling has a child. This contagion effect is strongest when substantial cost savings from coordination appear most likely and when the individual herself has little child-care experience and would thus rely heavily on siblings for information—further evidence consistent with peer effects.

What role do peer effects play in an individual's decisions of whether and when to have children? Fertility is a natural setting in which peer effects might arise, as raising children is arguably easier and more enjoyable when the associated experiences, information, and costs can be shared with others. Fertility trends in the U.S. and Europe during the 20<sup>th</sup> century have displayed large swings over time and place; it has been suggested that social multipliers arising from peer effects might explain such "excess variance" in fertility by amplifying the minimal variance in underlying determinants of fertility (Montgomery and Casterline 1996; Kohler 2000). A better understanding of how peer effect may influence fertility could be important to policy makers: fertility not only affects the solvency of many government-run pension programs but also influences the composition of the labor market, as childbirth appears to increase the labor-force participation of men (Lundberg and Rose 2000) and to decrease that of women (Angrist and Evans 1996).

Despite the potential importance of fertility peer effects, they are poorly understood in the current literature. That individuals tend to have similar fertility-related behavior could be caused by endogenous group formation or by an omitted determinant of childbearing that is positively correlated among peers. However, many papers have instead hastily inferred evidence of peer effects from such within-group correlations.<sup>1</sup>

---

<sup>1</sup> Most of the papers that use micro-data to exhibit fertility peer effects are based on data from developing countries and focus on the use of contraception. Madhavan, Adams, and Simon (2003) find that having more peers outside of the extended family increases the use of contraception in Mali. But as the data are cross-sectional in nature, there is little hope of accounting for endogenous selection into peer groups. Behrman, Kohler, and Watkins (2002) more directly address this problem, using data from rural Kenya. They use panel data and include individual fixed effects, relying on changes in individuals' peer groups. They find that those with a contraceptive user in their peer group are more likely to use contraception themselves. However, as they note, their identification requires that people randomly change their peers; if, instead, people's peer groups are in part a reflection of their changing values, then interpreting causation is not justified.

Indeed, social interactions in fertility are often assumed to exist in order to explain macro-level phenomena even though micro-level evidence showing that individuals actually respond to their peers is lacking.

In this paper, I use primarily individual-level panel data from the Panel Study of Income Dynamics (PSID) to examine whether individual fertility behavior is consistent with the existence of peer effects. (I also replicate my main results using a similar, though smaller data set, the National Longitudinal Survey of Youth.) I consider a particular group of peers that might be especially relevant in the fertility context: siblings. Individuals often enjoy considerable interactions with their nieces and nephews. Such interactions could have a profound effect on how they feel about starting or enlarging their own families. The effect need not be positive—a niece or nephew might temporarily satisfy one's desire for children or deplete the resources one's parents might have otherwise devoted to help raise a child. In other words, nieces and nephews could be either substitutes or complements for one's own children.

I find that the probability of having a child rises by about 0.11 percentage points (15 percent) in the 24 months after the birth of a niece or nephew. This effect does not appear to be merely a shift in timing but rather an increase in the expected number of children, as it is not offset by lower rates of childbearing later on. Thus, nieces and nephews appear to act as complements to one's own children, and childbearing seems to exhibit a contagion effect among siblings. This result is not driven by unobserved heterogeneity among families, as individual (and, by construction, family) fixed effects are included and thus only variation across time is exploited. Unobserved common shocks to siblings' fertility also do not appear to explain the results, as an individual's

fertility rises only in the period after his sibling has a child and not before. Thus, the results provide a possible microfoundation for the peer effects that have long been used to explain the large swings in aggregate birth rates in the developed world.

While the majority of the paper focuses on establishing the existence of fertility peer effects, I also explore some possible mechanisms by which these peer effects might operate. The existing theoretical literature has proposed several mechanisms by which peer effects might influence individual decisions, a few of which have directly addressed fertility: peers sharing information (Bongaarts and Watkins, 1996) and peers changing the costs and benefits of having a child (Dasgupta, 2000). My approach is to identify situations in which sibling effects are strongest and then to draw some provisional conclusions about which mechanisms are most consistent with these patterns. This exercise is admittedly subjective, as distinguishing among stories of peer effects can at times devolve into semantics—for example, classifying a peer effect as operating by lowering costs as opposed to raising benefits can depend on rather arbitrary definitions. Nonetheless, a consideration of which mechanisms are most likely responsible for a given result affords a framework around which to organize the analysis.

The most striking finding is that fertility rises dramatically (30%) for both men and women after their *sister* has a child—yet the birth of a child to a brother appears to have no effect on an individual’s fertility. I interpret this result as supporting a cost-side mechanism. Mothers are more likely than fathers to decide whether to share child-care advice, whether to include nieces and nephews in their own children’s activities, and whether to provide hand-me-downs to nieces and nephews—and may be more likely to provide such services to their own siblings than to siblings-in-law. Furthermore, the

sibling effect is especially strong for siblings who grew up in poorer families; cost-savings may be especially important for these individuals.

Sibling effects are especially strong among siblings who are close in age and siblings living in the same state. It is difficult to say whether these results support a cost-side or benefit-side explanation, as in either case the effect would likely increase with proximity. A final noteworthy result is that the increase in fertility in the period following the birth of a niece or nephew is especially strong for couples having their first child. I interpret this result as suggesting that information sharing may play a role in the peer effect, as couples that have no parenting experience seem most responsive to the birth of nieces and nephews.

The paper is organized as follows. The first section discusses how peer effects might operate among siblings in the context of childbearing, exploring several different potential mechanisms that have been identified in the peer-effects literature. The second section describes the data and empirical strategy used for identifying such peer effects. The third section presents the basic results, and focuses on separating true peer effects from the effects one would observe from common, unobserved shocks. The fourth section determines how well different theories of sibling effects fit the data by looking at interaction terms and examining estimates across different subpopulations. The final section concludes and offers directions for future work.

## **1. Peer effects and siblings' fertility**

Peer effects are usually modeled by treating an individual's costs and benefits of taking a given action as a function of his peers' characteristics or behaviors (e.g.,

Benabou 1993). For example, the costs of earning good grades may fall if one has a high-achieving roommate who can serve as a study partner or tutor (see Sacerdote 2001). Similarly, the stigma costs of having a child out of wedlock or being arrested may fall if one's friends are engaged in similar behavior (see Case and Katz 1991). Peer effects have been found to be almost exclusively positive, in that individuals tend to imitate their peers.

Peer effects could be especially relevant among siblings, who generally spend even more time together than close friends do. Moreover, the time siblings spend together may be more formative—psychological studies have found that siblings are better at teaching one another than are close friends.<sup>2</sup> Though older siblings may no longer spend as much time together as they did when they were children, emotions ranging from friendship and cooperation to competition and jealousy may still exist and could lead to strong peer effects among adult siblings.

There are many possible ways in which peer effects could operate in the realm of fertility decisions. First, individuals might learn about childbearing through their siblings, and be either encouraged or discouraged to have children themselves. If they also learn parenting skills and avoid mistakes their siblings made, the cost of childbearing falls and the effect of learning is to increase fertility.

That siblings learn from one another during childhood is well established. Pine (1995) studies the vocabulary formation of sibling pairs and finds that the younger child not only learns his first 50 words at a younger age but also picks up words and phrases in a manner consistent with a “higher degree of observational learning.” Perhaps more relevant to the current application, Rogers, Rowe, and Harris (1992) find that younger

---

<sup>2</sup> See Azmitia and Hesser (1993).

siblings are sexually active at a younger age, consistent with their learning from or imitating their older siblings. Of course, in the current application, birth order may matter less (a hypothesis I will test) as younger siblings can have children before their older siblings, but the basic idea of information as the driving force behind peer effects remains the same.

Second, having children around the same time could increase the benefits of childbirth if siblings enjoy sharing common experiences or if they believe their children might enjoy the company of a similarly-aged cousin. Thus, having a child could exhibit “network externalities,” in which the benefit of an activity rises with the number of participants (Katz and Shapiro 1985).<sup>3</sup> If increasing the gap between the age of one’s child and the age of a sibling’s child decreases the extent to which one can share common experiences with a sibling, then network externalities would encourage siblings to have children around the same time.

Similarly, coordinating the timing of childbearing with one’s siblings could lower the associated costs. Siblings can provide free advice, childcare, and even durable goods such as clothing and toys. The “first mover”—the sibling who has a child first—may be unable to refuse once such services are requested; similar free-riding among siblings has been observed with respect to caring for elderly parents (Konrad et. al. 1993). Furthermore, the extent to which one can benefit from a first-mover sibling may decrease with the time one waits to have children: the sibling’s memory fades, making her advice less valuable; her own children grow older and are less likely to provide companionship

---

<sup>3</sup> Katz and Shapiro first formalized the term, illustrating it with the then nascent example of computers (the more people buy a certain type of computer, the larger the variety of compatible software that will be available to use with it. Hamermesh (1999) provides a less technical example: leisure, which is usually enjoyed more when consumed together with one’s friends.

for a younger child; and the childcare products her children used that were once state-of-the-art become out-dated.

One of the main challenges in the peer-effects literature is identifying the mechanisms through which peer effects operate: do peer effects work through the emergence of group leaders, through individuals' need to conform, or by some other means? Unfortunately, data on most peer groups (e.g., a classroom of students, the employees in an office) do not allow researchers to observe the nature of the relationships among peers. Employee data, for example, would not provide information on an individual's popularity or leadership ability.

In a group of siblings, however, the age, sex, and geographic location of different siblings provide observable variation in the types of relationships peers experience. We can investigate whether older siblings generate more powerful peer effects than younger ones; whether being in close proximity to one another is necessary or whether peer effects can operate across great distances. Of course, siblings are a very special type of peer group, so these results should be applied to other contexts with care.

## **2. Data and Empirical Strategy**

The basic empirical exercise in this paper is to model how the costs and benefits of having a child evolve over time. To that end, I estimate the probability of having a child as a function of variables that might appear in a fertility cost-benefit analysis (e.g., age, marital status, information on past fertility such as number of children and how long ago the last child was born), as well as time-varying information on the childbearing of siblings.

## *Data*

Such an exercise requires panel data that include information on siblings. I primarily use the Panel Study of Income Dynamics (PSID), which has followed an original group of about 8000 families since 1968. The essential aspect of the PSID for my purposes is that it also follows the families created by the children of the original 1968 families, so that I have an evolving record of the grandchildren born to the original 1968 parents.

I include only those 1968 families for which all immediate members are included in the PSID. I find these observations by matching the 1968 mother's total fertility indicator with the number of children interviewed, so that if a mother had a child who had left the household by 1968, her family is excluded from my analysis. Similarly, if any of the children in the 1968 households have missing fertility data later in the survey, I also exclude the entire family. Defining an original 1968 PSID immediate family as a "dynasty," the complete fertility history (as of 2001) for each dynasty in my subsample is known.

The resulting subsample includes 823 dynasties, which produced 1817 children (about 2.1 per dynasty). These children become the adult siblings that I observe over time. These 1817 siblings produced 3666 children of their own (grandchildren to the original 1968 parents). See Figure 1 for a graphical depiction of the different generations. I expand each of the 1817 adult sibling observations by the number of childbearing months (which I define, following the medical literature, as months falling between an individual's 15<sup>th</sup> and 44<sup>th</sup> birthdays) covered by the PSID. On average, each

sibling has 300 person-month observations, as some individuals were not yet 44 in 2001, the most recent year for which I have fertility data.

Summary statistics appear in Table 1. The typical person-month in the data set corresponds to someone who is about 30 years old and who already has a child and between two and three nieces and nephews. Nearly two-thirds of the person-month observations correspond to individuals who have already been married at least once. The key dependent variable is the dummy variable “had a child in current month.” Not surprisingly, this is a very low-frequency variable, indicating that the average probability of having a child in a given month is about 0.0058. This value suggests that the total fertility of the typical PSID sibling is about 2.1 (0.0058 children per month \* 360 months per observation = 2.088 children per observation).

I also use the National Longitudinal Survey of Youth 1979 (NLSY), which tracks a group of 12,000 individuals who were teenagers in 1979. Though the NLSY often sampled more than one sibling in a household, they do not usually sample all siblings, a condition my analysis requires as I need a complete fertility record for each sibling. Thus, subjecting the NLSY to the same sampling cuts as the PSID recovers only 195 dynasties, 463 siblings, and 730 grandchildren. Summary statistics for the NLSY subsample appear in Table 1. As the NLSY began later than the PSID, many individuals were still in their childbearing years at the time they were last observed (2001), so the ratio of grandchildren produced to total siblings is lower. As the average age is about four years less than that of the PSID, marriage rates are substantially lower.

### *Empirical strategy*

The general aim of the econometric analysis in this paper is to estimate the probability of having a child in a given month as a function of variables relating to one's own fertility as well as those of one's siblings. First, however, I seek to estimate a "baseline" fertility profile that accounts for variation in childbearing costs that arise independently of the influence of siblings. To this end, I include individual fixed effects as well as age-in-years fixed effects. Thus, I completely control for the variation in fertility across time that arises purely from individual heterogeneity and the aging of each individual. Additionally, I control for each individual's current number of children with separate fixed effects and include dummy variables for whether a child was born to an individual in the last 12 months, 13-24 months ago, and 25-36 months ago. These variables are essentially mimicking the fertility profile of a person if she were to have children independent of the behavior of her siblings.

While this "baseline" fertility profile is interesting in its own right, the focus of this paper is to estimate the deviation from this baseline profile caused by changes in the fertility of an individual's siblings. An obvious measure to include is the total number of nieces and nephews born to an individual by a given month. As the discussion in the previous section highlighted the importance of timing, I also include dummy variables coded as one if a sibling had a child in the last 12 months, 23-24 months ago, and 25-36 months ago.

The basic regression model that I estimate is:

$$(1) \quad \Pr(Y_{jt}=1) = \beta_0 + \beta_1^*(\text{Sibling had a child in last 12 months})_{jt} + \beta_2^*(\text{Sibling had a child 13-24 months ago})_{jt} + \beta_3^*(\text{Sibling had a child 25-36 months ago})_{jt} +$$

$$\beta_i^*(\text{Number of nieces and nephews already born})_{jt} \\ + \gamma^*X_{jt} + I_j + A_t + \varepsilon_{jt}$$

where the sibling variables are coded as described above,  $Y_{jt}$  is a dummy variable coded as one if a child was born to individual  $j$  in month  $t$ ,  $X_{jt}$  is a vector of individual time-varying variables covering marriage and own fertility history,  $I_j$  is a vector of individual fixed effects, and  $A_t$  is a vector of age-in-years fixed effects. I use a conditional fixed-effects logit model to address the incidental parameters problem that plagues non-linear fixed-effects estimates—though using a linear probability model instead does not change the results.

### 3. Results

#### *Estimating the average size of the sibling fertility effect*

The first column of Table 2 displays the results from estimating equation (1) when the coefficients are reported as logit odds ratios. The coefficients on the variables relating to individual characteristics do not contain any surprises. There is a strong, negative effect of having had a child in the last 12 months, a positive effect of having had a child 13-24 months ago, and an even larger positive effect of having had a child 25-36 months ago. The patterns of these coefficients reflect the well-established tendency of parents in developed nations to space their children about two years apart (see Heckman-Walker 1990; Merrigan-St. Pierre 1998). There is also a large, positive coefficient on the marriage dummy variable.

The key variables of interest in this paper, however, are not those corresponding to own fertility, but to siblings' fertility. The coefficients on these variables are striking.

Individuals are 15 percent more likely to have a child in the 12 months following the birth of a niece or nephew and 17 percent more likely in the 13-24 months following the birth of a niece or nephew. The total number of nieces and nephews does not seem to play a significant role, however.

Column 2 shows the results when the same equation is estimated with a linear probability model instead of the conditional logit specification. Though the coefficients are not immediately comparable, dividing the percentage point changes reported in Column 2 by the baseline probability of 0.00586 yields values very close to those in Column 1. Finally, Column 3 reports the odds ratios from a conditional logit estimation of equation (1) when data from the NLSY are used. The pattern of the coefficients of interest is exactly the same as when the PSID is used: fertility rises in the first 12 months after a sibling has a child and in the 13-24 months after a sibling has a child (by 30% and 23%, respectively), but the effect 25-36 months out as well as the effect of the total number of nieces and nephews are insignificant.

#### *Peer effects or independent birth spacing?*

Do the results in Table 2 actually reflect sibling effects, or could siblings acting independently produce similar estimates? In particular, the reader might be wondering whether siblings have children that are spaced two years apart simply because they themselves are spaced two years apart. To make this concern more concrete, suppose the oldest sibling first gives birth at the age of 28, and, two years later, the next oldest sibling also has her first child at age 28. Suppose both siblings go on to have their second child at age 30. Then it appears that the younger sibling is reacting to the older one when in

fact she is following her own fertility pattern, which is the same as that of her older sister, just lagged two years behind.

However, in the specification represented by equation (1), the sibling effects described in Table 2 are *in addition* to any such coincidental correlations in the fertility patterns of individual siblings. The key point is that age measured in years is completely absorbed in the regressions, so the positive sign of the sibling coefficients cannot be attributed to fertility generally peaking at certain ages.

Not only is the complete age-fertility profile of each individual embedded in equation (1), but the individual's past fertility behavior is accounted for as well. The concern that individuals spacing their children two years apart is somehow generating the sibling effects is addressed by including dummy variables for whether an individual had a child born to her 1-12, 13-24, and 25-36 months ago.

Finally, one might be concerned that, even if the overall age-fertility profile is controlled for, different groups might peak at different ages. To address this concern, I interact each age-in-years dummy variable with a dummy variable for being female and a dummy variable for having finished high school and add these additional controls to equation (1). A comparison of Columns 2 and 4 reveals that the coefficients of interest are largely unchanged when the age-in-years fixed effects are interacted with sex and education variables. Allowing each of the age dummy variables to have a different effect for each of these groups is a fairly demanding request of the data and the fact that the coefficients remain largely unchanged should ease any concerns about the fragility of the estimates.<sup>4</sup>

---

<sup>4</sup> Note that I estimate the equation with the added fixed effects using a linear-probability model instead of a conditional logit model. Given the large number of extra parameters created by interacting each of the 30

Figure 1 illustrates the magnitude of the sibling effect when it is superimposed upon the baseline age-fertility profile of an individual. Predicted fertility, generated from the linear probability model estimated in column 2 of Table 2, tracks actual fertility fairly well. Not surprisingly, fertility peaks in the mid- to late-twenties. Figure 1 also shows the predicted effect of the birth of a niece or nephew, again using the results in column 2. Assuming said birth takes place at age 27, we see a sharp increase in fertility over the next three years. As fertility peaks near the age of 27, the sibling effect appears more muted than it would if placed during a period of low fertility, but is substantial nonetheless.

*Peer effects or independent responses to unobserved shocks?*

When researchers attempt to measure peer effects in cross-sectional data, they need to worry about unobserved heterogeneity among groups: after all, the “excess variance” approach used by Glaeser, Sacerdote, and Scheinkman (1996) is only valid if one has actually accounted for the variance in relevant control variables. The analogue in this setting is the need to control for heterogeneity in the incentives to have children *over time*. An unobserved shock to fertility that affects all siblings (e.g., a local hospital begins to offer infertility treatment or an adorable baby is born to a cousin or family friend) may look like a sibling peer effect even though siblings could be each acting independently.

How can peer effects be distinguished from independent reactions to unobserved shocks? Suppose that instead of merely considering the months after a sibling’s birth

---

age-in-year fixed effects with sex and education categories, estimation of this equation using conditional logit was not possible. Therefore, I used a linear probability model and comparisons should be made with the linear-probability estimate of the baseline equation in column 2.

occurs, dummy variables are added corresponding to the time *before* a birth occurs. If common shocks were driving the result, then these added coefficients should be no different than those tracking the months after a sibling has a child. To see this, imagine that some shock occurs: the siblings react, and some have children more quickly than others. The faster ones will raise the lead coefficients (their fertility is high in the months *before* their other siblings have a child) and the slower ones will raise the lag coefficients. On average, the coefficients should be symmetric around zero.

Column 5 shows the results from performing this exercise. While the coefficients corresponding to the period before a sibling gives birth are positive, they are close to zero and not statistically significant; the coefficients corresponding to the period after a sibling gives birth remain positive and significant and generally unchanged from the original specification. Thus, instead of exhibiting the symmetric pattern associated with the unobserved-shock story, fertility appears to respond only *after* a sibling gives birth, suggesting that individuals are reacting to their siblings—not responding independently to a common shock.

*Are peer effects being driven by siblings reacting to one another or planning together?*

Of course, one sibling reacting to another is not the only form peer effects can take; siblings can also explicitly plan together. Suppose that two siblings decide together to try to conceive with their respective spouses. Given the hazard function of time-to-pregnancy estimated in the medical literature, they conceive a child within six months of each other with probability of about 0.87.<sup>5</sup> Now, consider two siblings who did not plan

---

<sup>5</sup> This calculation is made by the author based on information contained in Gnath *et al* (2003), the largest prospective study to date of time-to-pregnancy. The authors calculate the probability of becoming pregnant

their pregnancies together. The first gets pregnant during month  $T$  and perhaps informs the second sibling during month  $T+3$ . After thinking it over for a while, the second sibling and her partner decide to try to have a child as well. If they were to conceive very quickly, they might be able bear a child within six months of the first sibling. Obviously, a sibling who waited until the actual birth of her niece or nephew before deciding she also wanted a child would generally be unable to have a child within six months of her sibling having done so. Thus, as a first approximation, increases in fertility within six months of a sibling's having a child could be due to either siblings explicitly planning together or siblings reacting to one another, but the increases in fertility beyond that point are probably due to siblings reacting to one another.

A more detailed look at fertility patterns surrounding the time of birth of a niece or nephew appears in Figure 3. Instead of dividing time into twelve-month intervals, I use six-month intervals, and plot the coefficients on these six-month dummy variables, focusing on the 24 months before and after the birth. As might be expected, given the probability of siblings having children within six months of each other when they begin trying at the same time, the six-months-before and six-months-after coefficients are large and positive. (The slightly larger coefficient for the six-months-after variable might reflect an individual who did not plan with her sibling but was able to conceive very

---

each month during a year in which couples are actively trying to conceive. Unfortunately, these data do not allow me to calculate perfectly the relevant statistic: the probability of two siblings having a child within six months of each other given that they begin to actively try to conceive at the same time. First, the probability calculated is actually based on the *joint* probability of two siblings conceiving with their respective partners within six months of each other *and* their both conceiving within 12 months of actively trying. Thus, the calculation I make will be a slight underestimate of the probability of two siblings having a child within 6 months of each other, as it would not include an event such as the two siblings conceiving 14 and 18 months after they began actively trying. Second, the Gnoth data refers to conception, whereas I observe births in the PSID data. However, this discrepancy would probably just add some noise to the estimate and not bias it greatly. Finally, as differences in sibling fecundity would be less than the differences between individuals chosen from the population at random, the chance of two siblings conceiving within 6 months of each other given that they start trying at the same time would be even higher.

quickly upon hearing the news of her sibling's impending birth.) However, any symmetry ends there. Fertility 12 and 18 months after a sibling has a child is markedly elevated, while fertility 12 and 18 months before a sibling has a child is not statistically different from zero.

If we take symmetry as a sign that siblings have coordinated their fertility decisions, or, as suggested earlier, as a sign that they are responding to a common shock, then these alternative hypotheses can explain elevated fertility only in the six months after a sibling has a child. Thus, it appears that the majority of the rise in fertility during the two years surrounding the birth of a niece or nephew comes from individuals reacting to their siblings.

#### **4. Different explanations of the sibling effect**

Though attempts to distinguish between siblings explicitly planning together and siblings reacting to each other provided some understanding of the mechanism driving the sibling fertility effects, it is still unclear why siblings do not make their fertility decisions independently of one another. By examining the differential strength of the fertility contagion effect, many of the explanations offered in Section 2 can be explored.<sup>6</sup>

---

<sup>6</sup> Some results are best viewed when the data are divided into different subsamples and coefficients are compared across regressions. Results of this type are reported in Table 3. Other results are clearer when the effect of a certain type of sibling (e.g., sisters, older siblings) can be measured relative to the effect of other siblings. Results of this type are reported in Table 4. I will be discussing the results thematically and not by table, which unfortunately requires the reader at times to flip between the two tables.

*Do siblings learn about childcare from one another?*

Siblings who have already had children might be the first source a couple turns to in order to find out what it's like to be a parent. Part of the sibling effects found so far could be the result of childless couples waiting until they can witness the experiences of a sibling. To test this possibility, I estimate the original regression in equation (1), but now include only those person-months corresponding to individuals who have already had at least one child. I report the results in column 2 of Table 3. Comparing these estimates to the original estimates, which are reproduced in the first column of Table 3, a couple of differences stand out. The effect of a sibling having had a child in the past twelve months is slightly larger for the subsample that has already had a child than for the entire sample. However, the situation is very different when the following twelve months are considered. There is a large, significant effect when the entire sample is considered, but the effect is far smaller and not statistically significant when those who do not yet have any children are removed. These results are consistent with childless siblings taking the time to extract information not only from a recent pregnancy of one of their siblings, but also the earliest phases of childcare. That childless couples do not react immediately to the birth of a niece or nephew is not surprising, as such behavior would allow for little information to be gained.<sup>7</sup>

As learning seems to be playing some part in the sibling effect, the effect of older siblings on younger siblings could be differentially strong as, at least in childhood, information generally flows in that direction. Indeed, the previously cited work on

---

<sup>7</sup> A natural approach would seem to be a comparison of a hazard estimation where the failure event is defined as the birth of a first child to that where the failure events is defined as the birth of a second child. Unfortunately, there is no exact analogue to a fixed-effect estimation in the hazard context and thus no way to estimate the effect of the sibling variables using only “within” variation, as is implicitly done with the fixed-effects conditional logit results.

childhood sibling effects suggests this asymmetric treatment pattern. Siblings may be unaccustomed to looking to younger siblings for guidance, or, alternatively, younger siblings may just blindly imitate older ones. To test this idea, I add the variables “an older sibling had a child in the last  $n$  months” with  $n$  taking the values 12, 13-24 and 25-36, and “total number of nieces and nephews born to older siblings by month  $t$ ” to equation (1). There are now eight fertility variables: the four original variables correspond to the average sibling effects, and the four “older sibling” variables represent the differential effect that older siblings have vis-à-vis younger siblings.

Column 1 of Table 4 shows that while the effect of older siblings is slightly stronger in the first 12 months, the effect of younger siblings is stronger in the following 24 months. There is also a small, negative effect of the total number of children born to younger siblings ( $0.957 < 1$ , though this estimate is not significant at the 0.05 level), and no absolute effect for the total born to older siblings ( $0.957 * 1.126 = 1.077$ ). Overall, it seems as though there is no consistently stronger effect of older siblings. Thus, while siblings do seem to learn from one another, information about having children seems just as valuable when it comes from a younger sibling as when it comes from an older one.

*Does sibling fertility display network externalities and/or free-ridership?*

Beyond the provision of information, peers can influence each other’s behavior when “network externalities” or free-rider opportunities are present. As discussed in Section 1, siblings might enjoy having children more if they can share their experiences with one another. Though a direct test of this hypothesis is difficult, it is reasonable to suppose that such network externalities should be stronger when peers are “closer.” I

present several empirical tests meant to determine whether siblings who appear “close” to one another exhibit stronger fertility contagion effects, where I define “close” in several different ways. Of course, determining that proximity increases peer effects does not tell us whether the effect is via the cost or benefit side; the claims I make below that some results support a cost interpretation while others a benefit interpretation are meant to be merely suggestive.

I first consider siblings who are close in age. Although the earlier analysis of older siblings did not reveal a hierarchical effect, there could be a symmetric effect. I define siblings as “close in age” if they are born within four years of each other (based on this definition, about half of all sibling pairs are “close in age”). I then create four additional “close-in-age” sibling variables: dummy variables indicating that a child was born to a close-in-age sibling 12 months ago, 13-24 months ago, and 25-36 months ago, and a running total of the number of children born to close-in-age siblings. Column 2 of Table 3 shows the results. As with the older-sibling analysis, the differential effect of close-in-age siblings relative to other siblings is represented by the coefficients on the “differential effect” variables; the absolute effect of close-in-age siblings is obtained by multiplying the main effect and the differential effect.

The results suggest that close-in-age siblings account for much of the observed sibling contagion effect and have a markedly different effect than siblings not close in age. Fertility rises by 39% ( $1.63 \times 0.86$ ) in the twelve months after a close-in-age sibling has a child; 29% ( $1.04 \times 1.24$ ) 13-24 months after; and 12% ( $0.988 \times 1.136$ ) 25-36 months after. The coefficients on the main effects show that children born to much older or much younger siblings have no positive effect on fertility, and in fact are associated with

a small decrease in fertility in the 12 months after such a sibling has a child. As siblings who are closer in age probably spent more time together as children and identify with each other more easily, these results suggest that, not surprisingly, peer effects grow more powerful as the intimacy between individuals grows stronger.

Peer effects might be affected by geographic proximity as well. For most observations in the sample, the individual's state of residence at the time of the interview is given. Thus, I have a rough sense of the location of each sibling in a given month. I consider two siblings as being "local" to one another at a point in time if they are living in the same state, and construct variables analogous to those used in the previous columns of Table 3 to pick up the differential effect of a "local" sibling having a child. As column 3 shows, the differential effect of local siblings is always positive and significant (except for the 13-24-months variable, which is also positive but just misses conventional significance with a p-value of 0.113). Interestingly, the total number of non-local siblings has a positive effect on long-term fertility, but the total number of local siblings does not (and the point estimate is actually negative, though not significant  $(1.17 \times 0.80 = 0.936 < 1)$ ).<sup>8</sup>

The last column in Table 3 examines whether one's sister's having a child has a different impact on one's fertility than one's brother having a child. While there is essentially no effect on fertility when a child is born to a brother, fertility increases by about 30% in the 36 months after a sister has a child. Although I do not show the results

---

<sup>8</sup> Unlike the other variables, which were extracted from the Individual File of the PSID, state-of-residence variables had to be merged from the Family File. As the merges were not perfect, I often have observations with missing state data. I have assumed that, when the location information for a person is missing for a given year, he has not moved since his last known state of residence. When individuals with missing information are dropped, the results do not change, though the standard errors are larger due to smaller sample size.

in the interests of space, these coefficients remain essentially unchanged when the same specification is estimated on only the women in the sample and then only the men in the sample. Thus, the birth of a child to a brother has no effect on his sisters or his brothers, but the birth of a child to a sister has an equally strong effect on both her sisters and brothers.

That sisters have a larger effect on sibling fertility could arise from the differential impact of childbirth on men and women. Childbirth may simply become more salient when one's blood relative physically experiences pregnancy and childbirth. When a child is born to a brother, the drama and excitement of the pregnancy and birth are centered around an in-law, perhaps diminishing the impact.

More economic explanations are also possible. The birth of a child might induce a woman to leave the labor force, giving her time to look after her own children and, perhaps, those of her siblings. Siblings (male or female) might expect or trust their sister, but not their sister-in-law, to look after their children. Therefore, the birth of brothers' children has less of an effect, even if their wives' labor-force participation also falls. As part of a mother's investment in her child may be expropriated by new parents in the family, the period immediately after the birth of a child to a female sibling may represent an especially fortuitous moment to have a child.

Similarly, the sister effect could be driven by the more extensive control mothers typically have over the raising of children, relative to fathers. Individuals might look to their siblings not only for free childcare, but also for free advice and information as well as durable goods such as baby clothes, toys, and strollers. A simple model in which individuals seek to aid the reproduction of blood relatives (a common finding in studies

of evolutionary biology) and in which mothers control the distribution of goods and services related to childcare would result in sisters helping their siblings raise their children. Brothers would want to help their siblings, but would have less authority to do so and, perhaps, less time as they are more likely to be fully employed after the birth of a child than would be the child's mother. Thus, the price of childbearing only falls when one's sister gives birth. Note that one might test this hypothesis by examining brothers married to only children, but information on the family background of spouses is not included in the survey.

Although these potential explanations for the sister effect focus on cost-side explanations, one could come up with benefit-side explanations as well. The proximity results (being close in age or living in the same state) could point to benefit-side stories, as one might enjoy more interactions with someone who lives nearby. On the other hand, one might also be able to extract more favors from someone close by, supporting a cost-side explanation.

One final check to see whether costs play a role is to separate the richer dynasties from the poorer dynasties. I divide the sample in half based on the household income of the original 1968 household (the parents of the siblings in the current analysis). The last column of Table 3 presents the results. The magnitude of the contagion effect is larger for the poorer families, and essentially disappears for the richer families: siblings from poor families increase their fertility by 15, 17, and 10% 12, 13-24, and 25-36 months after their sibling has a child, respectively. Among rich families, there is no statistically significant effect on fertility of a sibling having a child in the past 36 months. While this result does not in any way disprove peer effects operating on the benefit-side, it does

strengthen the interpretation that at least part of what is driving the effect is that the cost of childbirth falls in the immediate aftermath of the birth of a sibling's child. Assuming decreasing returns to wealth, rich families will be less responsive to any cost savings provided by nieces and nephews.

## **5. Discussion and conclusions**

One of the main goals of this paper has been to determine, using micro-level data, whether the peer effects so often used in models of aggregate birth rates actually exist. Examining the fertility patterns of siblings, I find a large contagion effect: in the 24 months after an individual has a child, the probability that her siblings also have children increases significantly. The strength of the contagion effect varies across subpopulations in a manner consistent with the mechanisms proposed in the theoretical literature: the sibling fertility effect appears to be driven in part by information and in part by cost-sharing or network effects.

Obviously, individuals have many peers in addition to their siblings. An interesting extension might be to track groups of friends to determine whether similar contagion effects exist. I know of no publicly accessible dataset adequate for this task, but datasets could be created with this goal in mind (using retrospective questions would avoid the need for extensive follow-up surveys). The current paper thus provides a lower bound for the true estimate of the fertility peer effect, as only a small subset of one's peers are observed. Furthermore, estimating the full peer effect would allow those modeling aggregate fertility trends using peer effects to calibrate their models instead of being forced to choose the strength of the peer effect rather arbitrarily.

Finally, the sibling effect might help to explain the apparent “inheritance” of family size: individuals from large families tend to have many children. A favorite topic of the first modern statisticians,<sup>9</sup> the reasons behind the intergenerational fertility correlation have been debated since the end of the 19<sup>th</sup> century: does the correlation reflect inherited differences in biological fecundity, inherited differences in preferences, or is it merely an artifact of the omission of relevant environmental variables in the childbearing function? As biological research has cast doubt on the fecundity theory, the literature has generally converged on some vague explanation of preferences for family size being passed on to the next generation.

The sibling effect suggests that fertility might be correlated between generations even when fecundity and preferences are not. Assuming that individuals from large families and small families are identical with respect to their ability and desire to have children, then individuals from large families will still have more children because they more often experience the positive “shock” to fertility due to a sibling having a child, simply because they have more siblings. Such a model would imply that fertility shocks to a dynasty will persist over time. The application of sibling effects to intergenerational models of fertility appears to have many testable predictions and could be a promising area for future work.

---

<sup>9</sup> See Pearson and Lee (1899), for the earliest statistical treatment of the phenomenon.

## References

- Angrist, J. and W. Evans, "Children and their Parents' Labor Supply: Evidence From Exogenous Variation in Family Size," NBER Working Paper #5778 (1996).
- Azmita, Margarita and Joanne Hesser, "Why Siblings are Important Agents of Cognitive Development: A Comparison of Siblings and Peers," *Child Development*, April 1993, Vol. 64, No.2, 430-444
- Benabou, R. "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics*, Vol. 108, No. 3 (August 1993).
- Butcher, Kristin and Anne Case. "The Effect of Sibling Sex Composition of Women's Education and Earnings," *Quarterly Journal of Economics*, August 1994, v. 109, iss. 3, pp. 531-63
- Chamberlain, Gary. "Multivariate Regression Models For Panel Data," *Journal of Econometrics*, 1982, v18(1), 5-46.
- Checkovich, Tennille J; Stern, Steven. "Shared Caregiving Responsibilities of Adult Siblings with Elderly Parents," *Journal of Human Resources*. Vol. 37 (3). p 441-78. Summer 2002.
- Raquel Fernandez and Fogli, Alessandra. "Fertility: The Role of Culture and Family Experience," NBER WP #11569, Aug 2005
- Garces, E, D. Thomas and J. Currie. "Longer-Term Effects of Head Start," *American Economic Review*. Vol. 92 (4). p 999-1012. September 2002.
- Goldin, C. and L. Katz, "The Power of the Pill," *Journal of Political Economy*, Vol. 110, No. 4 (2002).
- Heckman, James J; Walker, James R. "The Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data," *Econometrica*. Vol. 58 (6). p 1411-41. November 1990.
- Jensen, T. et al, "Regional differences in waiting time to pregnancy among fertile couples from four European cities," *Human Reproduction*, Vol. 16, No. 12, pp. 2697-2704 (Dec. 2001).
- Keiding, N. et al, "Estimating time to pregnancy from current durations in a cross-sectional sample," *Biostatistics*, Vol. 3, No. 4 (2002), pp. 565-578.

Konrad, Kai A, et al. "Geography of the Family," *American Economic Review*. Vol. 92 (4). p 981-98. September 2002.

Lundberg, Shelly; Rose, Elaina. "Parenthood and the Earnings of Married Men and Women," *Labour Economics*. Vol. 7 (6). p 689-710. November 2000.

Merrigan, Philip; St-Pierre, Yvan. "An Econometric and Neoclassical Analysis of the Timing and Spacing of Births in Canada from 1950 to 1990," *Journal of Population Economics*. Vol. 11 (1). p 29-51. February 1998

Murphy, M; Knudsen, L B. "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*. Vol. 56 (3). p 235-48. November 2002

Oettinger, Gerald S. "Sibling Similarity in High School Graduation Outcomes: Causal Interdependency or Unobserved Heterogeneity?" *Southern Economic Journal*. Vol. 66 (3). p 631-48. January 2000.

Parish, William and Robert Willis, "Daughters, Education, and Family Budgets: Taiwan Experiences," *Journal of Human Resources*, Fall 1993, 28(4), pp 862-98

Sacerdote, Bruce. "Peer Effects With Random Assignment: Results For Dartmouth Roommates," *Quarterly Journal of Economics*, 2001, v116(2,May), 681-704

Scheike, Thomas and Tina Jensen, "A Discrete Survival Model with Random Effects: An Application to Time to Pregnancy," *Biometrics*, Vol. 53, No. 1. (Mar., 1997), pp. 318-329.

Wooldridge, *Econometric Analysis of Cross-Sectiona and Panel Data*. MIT Press, Cambridge, MA (2002).

Table 1: Summary statistics

	<i>Mean</i>	<i>Stand. dev.</i>	<i>Mean</i>	<i>Stand. dev.</i>
Had a child born in current month	0.006	0.076	0.005	0.073
Female	0.519	0.500	0.481	0.500
Total number of children born by current month	1.248	1.106	0.752	1.070
Married at least once by current month	0.635	0.481	0.483	0.500
Tot, no. of nephews and nieces born by current month	2.898	3.703	1.154	1.631
Age in current month	29.953	9.154	328.586	85.974
Data source	PSID		NLSY	
Observations	625,151		135,491	

Notes: To be included in this table and the following tables, an individual must have been one of the original children interviewed in the 1968 PSID (1979 NLSY), must be a child of one of the original adults interviewed, and must have no missing fertility information. Additionally, all siblings of the individual must also meet these criteria.

Table 2: The effect of sibling fertility on own fertility

	<i>PSID, logit odds ratios</i>	<i>PSID, linear probability</i>	<i>NLSY, logit odds ratios</i>	<i>PSID, linear probability</i>	<i>PSID, logit odds ratios</i>
Sibling had a child in past 12 months	1.149 (0.060)	0.00102 (0.00037)	1.296 (0.164)	0.000844 (0.00037)	1.148 (0.060)
Sibling had a child 13-24 months ago	1.167 (0.061)	0.00116 (0.00036)	1.226 (0.153)	0.000970 (0.00036)	1.164 (0.061)
Sibling had a child 25-36 months ago	1.062 (0.056)	0.00047 (0.00035)	0.900 (0.120)	0.000265 (0.00035)	1.059 (0.056)
Sibling will have a child within 12 mos					1.040 (0.056)
Sibling will have a child within 13-24 mos					1.054 (0.056)
Sibling will have a child within 25-36 mos					1.044 (0.056)
Total number of nieces and nephews	1.002 (0.018)	-0.00003 (0.00009)	1.030 (0.065)	-0.000041 (0.00009)	1.012 (0.020)
Have been married at least once	7.668 (0.586)	0.01774 (0.00063)	10.239 (1.881)	0.016898 (0.00064)	7.648 (0.585)
Had a child within 12 months	0.417 (0.062)	-0.00600 (0.00034)	0.000 (0.000)	-0.006123 (0.00035)	0.417 (0.062)
Had a child 13-24 months ago	3.087 (0.238)	0.00399 (0.00052)	1.848 (0.312)	0.003757 (0.00053)	3.088 (0.238)
Had a child 25-36 months ago	3.725 (0.262)	0.00660 (0.00053)	3.113 (0.465)	0.006292 (0.00053)	3.724 (0.262)
Type of Fixed-effects	Age (in yrs)	Age (in yrs)	Age (in yrs)	Sex*age, ed*age, age	Age (in yrs)
Observations	615,641	625,151	98,732	625,151	615,641

Notes: The level of analysis is the person-month. Columns 1, 3 and 5 report results as odds-ratios while other columns report results as changed in probability. All regressions include individual fixed effects, in addition to whatever fixed effects are specified in the table. Columns 1, 3, and 5 employ a conditional logit specification in order to avoid the incidental parameters problem.

Table 3: Effect of sibling fertility on own fertility across different subpopulations

	<i>All observations</i>	<i>Those with a child already</i>	<i>Below median income</i>	<i>Above median income</i>
Sibling had a child in past 12 months	1.149 (0.060)	1.201 (0.093)	1.145 (0.082)	1.094 (0.085)
Sibling had a child 13-24 months ago	1.167 (0.061)	1.064 (0.084)	1.172 (0.085)	1.096 (0.086)
Sibling had a child 25-36 months ago	1.062 (0.056)	1.077 (0.081)	1.099 (0.080)	0.986 (0.078)
Total number of nieces and nephews	1.002 (0.018)	1.013 (0.032)	0.991 (0.023)	1.060 (0.036)
Total observations	615641	319634	303586	291563

Notes: See table 2.

Table 4: Effect of different types of siblings' fertility on own fertility

	<i>Type: Sibling is older than I</i>	<i>Type: Sib is close in age to me</i>	<i>Type: Sibling lives in my state</i>	<i>Type: Sibling is my sister</i>
A sibling had a child in past 12 months	1.059 (0.076)	0.858 (0.064)	0.826 (0.070)	1.002 (0.074)
A siblings had a child 13-24 months ago	1.247 (0.094)	1.036 (0.081)	1.023 (0.098)	0.990 (0.080)
A sibling had a child 25-36 months ago	1.134 (0.093)	0.988 (0.078)	0.906 (0.092)	0.893 (0.074)
Total number of nieces and nephews	0.957 (0.023)	0.986 (0.028)	1.174 (0.039)	0.974 (0.030)
This type of sibling had a child in past 12 months	1.182 (0.106)	1.636 (0.143)	1.586 (0.147)	1.255 (0.110)
This type of sibling had a child 13-24 months ago	0.907 (0.088)	1.242 (0.120)	1.187 (0.129)	1.307 (0.127)
This type of sibling had a child 25-36 months ago	0.911 (0.092)	1.136 (0.111)	1.235 (0.141)	1.313 (0.131)
Tot. no. of children born to siblings of this type	1.126 (0.044)	1.033 (0.046)	0.803 (0.033)	1.053 (0.046)
Observations	615641	615641	615641	615641

Notes: See Table 2.

**Figure 1:** Three generations of the PSID

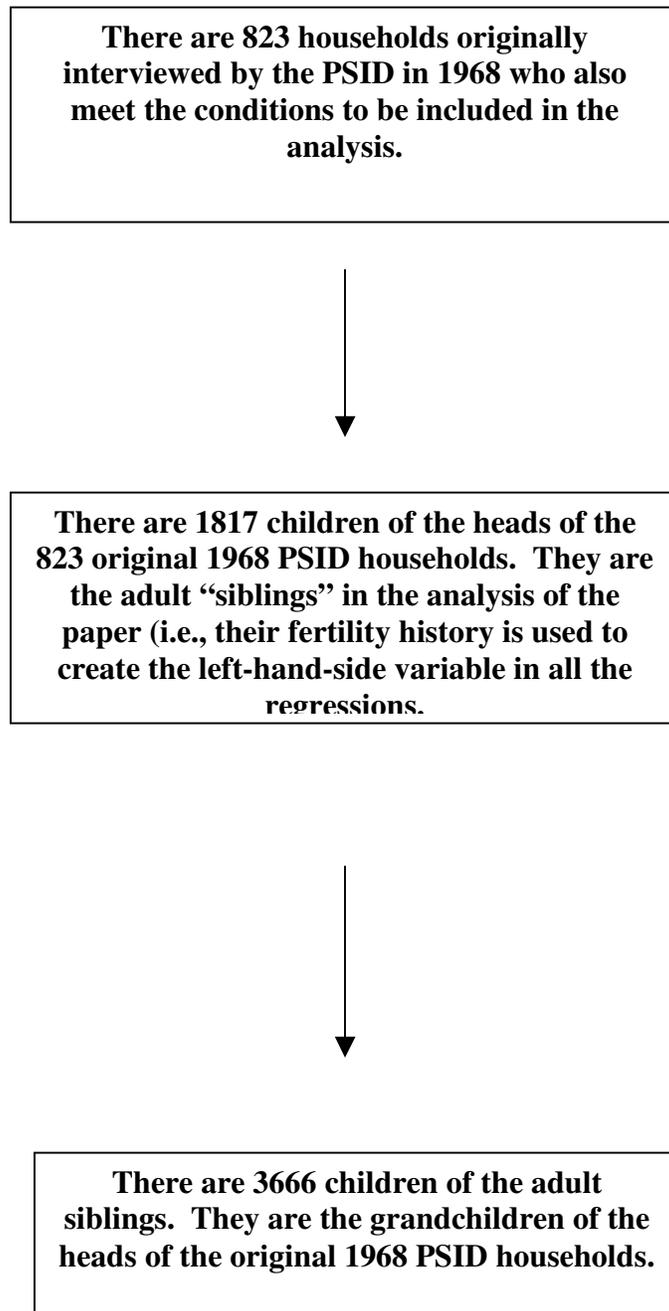


Figure 2: Age-fertility profile

