

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

EUI Working Paper **ECO** No. 2001/5

Fertility and Divorce

DANIELA VURI

BADIA FIESOLANA, SAN DOMENICO (FI)

All rights reserved.

No part of this paper may be reproduced in any form
without permission of the author.

©2001 Daniela Vuri
Printed in Italy in March 2001
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

Fertility and Divorce

Daniela Vuri^{*†}

February 2001

Abstract

In this paper a model which relates a couple's divorce probability and fertility decision is developed. Theory predicts that couples with children are less likely to divorce, and conversely that couples with higher *ex-ante* divorce probabilities are less likely to give birth to children. The model's predictions are tested using the five waves 1990-95 of the German Socio-Economic Panel and of the Panel Study of Income Dynamics, and the five waves 1991-1996 of the British Household Panel Survey. The identification and estimation of the causal effect of fertility on divorce is based on instrumental variable estimation. The sex of the two previous children is chosen as an instrument for exogenous fertility movements. IV estimation results contrast strongly with simple OLS estimates. Once the problem of the endogeneity of fertility is taken explicitly into account, the implied instrumental variable estimate of the effect of fertility on divorce has positive and therefore opposite sign with respect to the conventional least squares estimates.

Keywords: Marital dissolution; Fertility; IV Estimation

JEL-Code: C2; C3; J1; J12; J13

*Mail address: Department of Economics, European University Institute, Via dei Roccettini 9, I-50016 San Domenico di Fiesole (FI). E-mail: vuri@iue.it.

†I would like to thank Andrea Ichino for comments, patience and guidance, Joshua Angrist, Guido Imbens, Massimo Motta and Daniela Del Boca for very helpful comments and advice, participants in the ESPE Conference and seminar participants at the EUI, especially Sascha Becker and Anna Sanz De Galdeano. All errors are mine.

1 Introduction

One of the most striking trends in postwar social indicators in the developed countries has been the rise in divorce rates. Table 1 illustrates the rate of divorce per 1000 persons in the Western countries over time. After staying at low levels for many years, divorce rates began to rise in the mid-1960s. In less than 40 years the number of divorces has doubled in the USA (from 2.2 in 1960 to 4.4 in 1995), and has augmented even more in Europe (e.g. in the UK the rate of divorce rose from 0.5 in 1960 to 2.9 in 1995, in Germany it increased from 1.0 in 1960 to 2.1 in 1995).

This “breakdown of the traditional family” has attracted much concern because, on average, divorce is associated with a deterioration of the economic status of women and children. Most studies¹ show that children in single parent families might suffer from the lack of parental investment, caused not only by the lower input from the absent parent, but also by the employment of the present parent.² Therefore, children fortunate enough to grow up within surviving marriages receive on average higher investment than do children whose parents never marry or marry and then divorce. Noting this, a society might attempt to improve the *lots* of its children not only by developing an institution of marriage, but also “strengthening” it. By penalizing divorce, for example, a society might try to reduce the incidence of “broken” marriages and thus increase the average welfare of children. However, “social penalties are crude instruments and their imposition creates a trade-off, because some mothers and children might benefit from the investments by fathers, but others suffer from being “trapped” in bad marriages” (Murphy, 1999).³

In response to the important consequences of marital dissolution and the increasing divorce rates in the last decades, researchers have

¹See Duncan and Hoffman (1985, 1988), Weitzman (1985), and McLahanan and Sandefur (1994) for an overview.

²See Ermish and Francesconi (1997) for a survey of the literature.

³Penalties could consist for example in separation periods with intensive counseling which would make divorce more costly (Gruber, 2000). However, I do not enter here into the questions of how social penalties might be imposed and what form they might take.

begun to investigate the causes of divorce. The most critical factors historically associated with the rise in divorce are the fertility decline and the increase in women's labor force participation.⁴

In this paper, I ignore the issue of female labor supply, already widely investigated by economists;⁵ instead, I focus on the relationship between fertility and divorce.

The apparent simultaneous decline in fertility and increase in divorce rates observed in the last decades leads naturally to the following question: can we infer from these figures that the presence of children discourages marital dissolution, that is, is there a causal nexus between the two? Or is there rather a problem of omitted variables bias, due to the presence of confounding factors that may jointly determine fertility and marital dissolution?

The objective of the present paper is to quantify the causal influence of children on marital (in)stability, setting up a theoretical framework that links fertility and divorce decisions and estimating this model using 2-stage instrumental variable techniques.

The remainder of the paper is organized as follows: Section 2 briefly summarizes the literature on fertility and marital instability; Section 3 formulates a dynamic decision model of fertility and divorce. Section 4 indicates how to move from the model to the data. Section 5 describes the methodology to be used in testing the potential problem of endogeneity of fertility; a brief description of the instrument chosen is also provided. Section 6 describes the data sets used (the German Socio-Economic Panel-GSOEP, the British Household Panel Survey-BHPS, and the Panel Study of Income Dynamics-PSID), and the process of sample selection. It provides some summary statistics as well. Section 7 contains the estimation results and evaluates the robustness of the results. Finally, Section 8 presents some concluding remarks and directions for

⁴In the USA, the fertility rate decreased from 3.65 in 1960 to 2.02 in 1995, in the UK from 2.72 in 1960 to 1.70 in 1995, and in Germany from 2.37 in 1960 to 1.25 in 1995.

⁵See Angrist and Evans (1998), Carrasco (1998), and Nakamura and Nakamura (1992) for an excellent review of the argument.

further research.

2 Review of the literature

Following Becker's (1974) pioneering work, a substantial body of literature on the economics of marital dissolution has begun to accumulate. The literature on the effect of children on marital disruption can be divided into two categories based on the methodological approach: studies considering fertility as an exogenous variable, and studies addressing the problem of endogeneity of fertility.

Studies in the first category do not offer any consensus on the effect of fertility on marital stability. This lack of consensus is due largely to the complex nature of the problem: the effect of children on marital stability is the result of the interactions of different factors (presence, numbers and ages of children), and the findings depend on which part of the problem is examined. Consequently, children can either promote or weaken marital stability. They can give stability to the marriage either by rising the cost of marital dissolution or by increasing the benefits derived from the marriage. For instance, couples who are unhappily married but are aware of the potential negative consequences of divorce on their children may avoid marital disruption.

White and Lillard (1991) find that firstborn children increase the stability of their parents' marriage through out their preschool years. Children after the first one decrease the chances of dissolution but only when they are very young, while children born before marriage increase significantly the chances that the couple will dissolve. These results are also reported by Becker, Landes and Michael (1977), Andersson (1997), Peter (1986), Ono (1998), and many others. From these results it could be argued that the motivation to avoid marital dissolution consists only in delaying it, since the effect seems to be relevant for very young children, who might suffer more from parent's divorce. Weiss and Willis (1997) propose an extensive model of decision making related to the variable of interest, marriage status, but they have left the investment

decision process untreated. As investment is not their primary concern, the model assumes that investment in children is exogenous and their estimation results show that the number of children has a significant and positive influence on marital duration. Lillard and White (1993), White, Booth, and Edwards (1986), and White, Haggstrom, and Kanouse (1985) estimate the impact of first birth on disruption within the two years following the birth. They report substantially lower disruption rates in comparison to the group of childless married couples. The higher probabilities of divorce for childless couples may be attributed to the absence of child-related costs.

Children may also encourage marital stability by increasing the benefits of marriage, that is, children may create satisfaction inside the marriage. Still, this does not seem to hold in the case of higher number of children (Thornton, 1977).

On the other side, children may decrease the benefits of marriage. The events surrounding the birth of children may produce distress in the family, and this is particularly true for families with a *large* number of children (Thornton, 1977). Furthermore, Jensen and Smith (1990) find no evidence for Denmark neither of a positive effect of the number of children nor of the number of children in different age-groups on marital stability. The explanation given by the authors to this puzzle is that, if both members of a married couple work, the presence of children may strengthen the conflicts about the division of labor inside the household, thus counterbalancing the positive effect of having children on marital stability. This result might be expected in a country like Denmark, where there is a high female participation in the labor market.⁶

However, if fertility is not an exogenous variable in the divorce equation, all these studies provide biased estimates of the effect of fertility on divorce.

Due to this problem, the second type of study discusses explicitly the problem of endogeneity of fertility: not only the presence of children

⁶In 1993, for example, the female employment rate in Denmark was 68.7, the highest in Europe.

affects marital instability, but also the *potential* stability of the parents' marriage may affect the arrival of children because they represent the most important commitment to marriage; consequently, couples with a relatively high probability of dissolution may delay in making this commitment (Peters, 1986).

Becker et al. (1977) address the problem of endogeneity of fertility. They build a sample restricted to white women aged 40-55 with their first marriage intact, whose fertility is already completed, in the hope of controlling for any effect of fertility on divorce.⁷ They find evidence that couples whose members are heterogeneous in age, religion or education (factors which should be indicative of the probability of divorce) tend to give birth to a lower number of children than couples more homogeneous in these characteristics. However, this analysis is limited: the procedure used is valid only if we accept the hypothesis made by the authors that the probability of divorce is captured by the differences in the observed couple's above-mentioned characteristics.

Koo and Janowitz (1983) formulate, for married couples, a simultaneous model of the interrelationship of the probability of separation and of having a birth during this period (when dissolution presumably is being considered). The results indicate that childbearing patterns - number of children and age of the youngest child at the beginning of the marital interval being studied and fertility during the interval - do not influence the likelihood of separation over the marital life course, nor does marital strife (as indicated by separation) seem to affect childbearing throughout the marriage.

Lillard and White (1993) test the hypotheses that a couple's risk of marital disruption affects the timing of marital conceptions and that the risk of marital dissolution is affected in turn by the presence and number of children born to the couple over time. To test the simultaneous relationship between marital dissolution and marital fertility, they

⁷Initially, they suggest the use of a simultaneous equations model to identify the causation between children and dissolution, but then they decide against this strategy. Instead, they attempt to study causation by constructing a situation, such as the one described above, that largely excludes reverse causation.

use a model which includes the probability of disruption as a predictor of timing and likelihood of marital conception, and then they include the results of previous fertility decisions as a predictor of the dissolution of marriage.⁸ They find evidence that the probability of disruption has strong negative effects on the probability of marital childbearing, decreasing the chances that a child will be born. However, there is no significant evidence of the opposite effect, running from fertility to the probability of divorce.

In a recent publication, Brien, Lillard and Stern (1999) propose a way to model endogenous investment in a model of cohabitation, marriage and divorce; unfortunately they revert to exogenous investment as an element of the cost of divorce in their estimation because of computational costs. Their results indicate that investment in children is a significant negative correlate of divorce probability.

This paper, while similar to those just mentioned above in that it takes into account the problem of endogeneity of fertility, differs from them in two ways.

Firstly, it formalizes the theoretical argument of the existence of a link between divorce and childbearing by building a two-period dynamic model of marital status and fertility decision. Secondly, it also investigates empirically the causal link running from fertility to the divorce decision by using the instrumental variables (IV) technique based on the sibling sex composition of the first two children.⁹

⁸The identification of the parameters of the two equations requires the placing of some restrictions. In particular, in order to identify the parameters of the fertility equation, they include measures of the educational costs in the current state of residence in this equation, but not in the dissolution equation. Then, in order to identify the parameters of the divorce equation, they include indicators of the legal environment for divorce in the state of current residence in the divorce equation, but not in the fertility equation.

⁹The sex-preference instruments were used for the first time by Maria Iacovou (1996) and Angrist and Evans (1998) to estimate the effect of fertility on female labour supply respectively in the UK and in the USA.

3 Theoretical framework

A useful tool for examining the relationship between fertility and divorce is Becker's analysis of marriage (1974), according to which marriages and cohabitation are seen as voluntary arrangements between two adults, formed to coordinate consumption and production activities, including the conception of children. In Becker's framework, persons marry when the expected utility from marriage exceeds the expected utility from remaining single. The utility from marriage is increased by the presence, even indirectly, of children. However, at the time of marriage, both partners have limited information on the mate and the gains of the marriage, and, later in the marriage, a divorce may occur.¹⁰

The economic theory of divorce suggests two general causes for marital dissolution.¹¹ First, the search for a partner is costly and meeting occurs randomly; thus, a union which is currently acceptable may become unacceptable if either partner meets a person who might be a superior match.¹² Second, traits which influence the benefits of a union can change over time in an unpredictable manner; such surprises can cause either of the partners to reconsider their original decision.¹³

However, even if the spouses find out, *ex post*, that they are not very well matched, they may have few incentives to separate if they have made a large number of investments "specific" to the marriage.¹⁴ The

¹⁰See Weiss (1997) for an overview of economic theories on marriage.

¹¹See Becker et al. (1977) and Becker (1981).

¹²For example, persons marrying much younger than average have significantly higher probability of dissolution because they are likely to have searched for a shorter time; therefore, they are more likely to make a poor match and to dissolve the union in the future.

¹³For example, an unexpected low or high income or unexpected spells of unemployment can affect partner dissolution.

¹⁴Instead of the "*marital-specific capital accumulation explanation*" offered by the economists, sociologists have offered a slightly different theoretical argument to explain the changes in marital disruption rates with the addition of children. Following Durkheim's argument that "the sexual division of labor is the source of conjugal solidarity", sociologists have maintained that childbearing and rearing produce greater role differentiation and, thus, greater interdependencies between wives and husbands.

investments made by a couple could be divided into two groups: the ones that are almost as valuable if their marriage dissolves because they retain their value regardless of the couple's marital status (e.g. automobiles). The others are "particular investments" because they are less valuable if the marriage dissolves, and are therefore called "marital-specific" (e.g. information on the partner's preferences, children, and housing). One immediate implication of this distinction between marital investments is the way in which they affect a couple's divorce probability because "the accumulation of marital-specific capital raises the expected gain from remaining married" (Becker et al. 1977, pag.1152).

Therefore, parenthood provides an important basis for marital stability and children greatly lower the risk of marital disruption because they represent the most important marital-specific investment of a couple during their marriage.¹⁵ The presence of children may also delay divorce by making it more costly than continuation in the marriage because of the anticipated complications attending a divorce action, such as problems with child custody, visitation plans, coparenting and single-parent problems. Furthermore, the increasing awareness of the financial and psychological costs of divorce for children may give some parents pause, and lead to delay.¹⁶ Consequently, *children appear to constitute financial, legal, and emotional¹⁷ barriers to divorce.*

However, the decision to have children depends critically on the perceived durability of the marriage. Couples who face a relatively high likelihood that they will not stay together may delay the decision to have children, because of the higher costs of ending a marriage with children with respect to one without (Weiss and Willis, 1985).

¹⁵See Becker et al. (1977), Cherlin (1977), Becker (1991), Morgan and Rindfuss (1985).

¹⁶According to Weiss and Willis (1985), a separation can lead to inefficiently low levels of child care. For instance, when a child's parents split, the absent parent (usually the father) is encouraged to spend less on his children because it is difficult to monitor how the custodial parent (usually the mother) will spend the money.

¹⁷*We are referring in this case to a sort of "stigma" that is sometimes attached to persons who divorce when they have children, especially very young, which might discourage couples from divorcing.*

In what follows, a two-period model of marriage status and fertility decisions which considers both the directions of causality is presented, drawing on Sullivan (1995). In particular, in the context of a model of marital-specific investment, it is shown that marriage continuation probability increases as the number of children increases, and that the number of children is increasing in an unobservable measure of the quality of the marriage, which in turn influences the perceived marriage duration.

3.1 A dynamic decision model of fertility and divorce

Consider a married couple that lives two periods, $t = 0, 1$. During the first period the couple decides how many children C to have. Assume that *the fertility decision is taken only in the first period*, and that it cannot be changed in the second period. Furthermore, assume that children provide *differing levels of marginal utility* depending on the couple's second period marital status. Then, in the second period, the couple decides whether to separate or not.¹⁸

Suppose that a match value of the marriage exists, which can also be interpreted as a stochastic gain to marriage, $\theta_1 \in (-\infty, +\infty)$.¹⁹ It has a density function $f(\theta_1)$ and is not revealed until after the investment decision is made. The couple has, however, some information on the match value θ_1 , because the couple's demographic and social characteristics have been known since the time of marriage.²⁰ Although this

¹⁸Instead of considering the couple as the “decision making” unit, the model could be reformulated in another way by specifying each marriage as having two agents who optimize separate utility functions. This allows for strategic interaction in investment and divorce decisions, as well as differing utility levels outside of marriage for the two agents. Both a non-cooperative and a cooperative mode of interaction could be analyzed for paired agents. This topic will be explored in greater depth in future research.

¹⁹ θ_1 could be thought as a gain to the division of labor within the household that may be particular to marriage or a lower price associated with rising children within marriage (see Weiss, 1994).

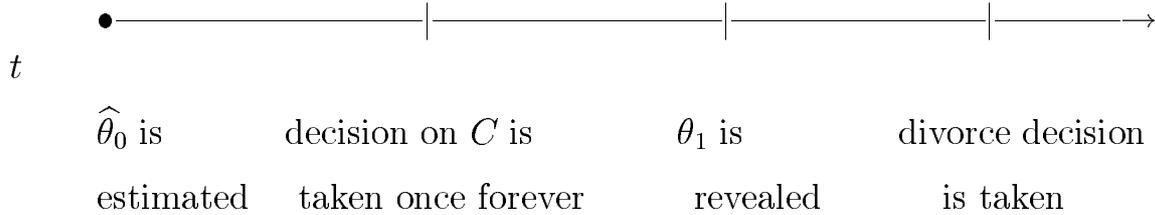
²⁰In the empirical part, section 5, we will divide explicitly the part of θ_1 which is un-

information allows the couple to make some predictions regarding its match value, some components of the match remain unobserved until later in the marriage. To reflect this in the model, I assume that the couple observes at time θ a noisy signal of the true quality of the match, $\hat{\theta}_0$, that satisfies the following property:

$$\frac{\partial \Pr [\theta_1 > c | \hat{\theta}_0]}{\partial \hat{\theta}_0} > 0 \quad (1)$$

This means that higher values of $\hat{\theta}_0$ act to right-shift the couple's subjective density of θ_1 . We can let $f(\theta_1/\hat{\theta}_0)$ be the couple's subjective density for θ_1 , given the noisy estimate $\hat{\theta}_0$.

Therefore, the timing of the decisions is as follows:



In period θ , the couple does not know the real probability of divorce in the next period (*ex-post* probability of divorce), but it can formulate an estimate of such a probability, the *ex-ante* probability of divorce, depending on the subjective estimate of the match quality $\hat{\theta}_0$. Once $\hat{\theta}_0$ is estimated, the couple decides if and how many children to have on the basis of the benefits and the costs implied by this decision. Then, in period 1 , the *real* match value θ_1 is revealed and the only choice the couple can make is whether to continue its marriage or to divorce.

Let's define the utilities for the couple of being married at time θ , of still being married at time 1 , and of divorcing at time 1 :²¹

observable, ϵ_i , and the part which is observable, given by the personal characteristics X_i .

²¹In equations (2), (3) and (4), I have assumed that the utilities depend only on children (eq. (3) also on the quality of the match). This could seem a strong assump-

$$U_0^M = m_0(C) \quad (2)$$

$$U_1^M = m_1(C) + \theta_1 \quad (3)$$

$$U_1^D = d_1(C) \quad (4)$$

where:

- i. U^i is the utility under marital status i ;
- ii. $m_0(C)$ values the net utility provided by children in period 0 ;
- iii. $m_1(C)$ and $d_1(C)$ value the net utility provided by children at time 1 under the two marital statuses;²²
- iv. it is reasonable to assume that children give different utilities through the couple's life cycle; hence, I allow $m_0(C)$ to be different from $m_1(C)$;²³
- v. I also assume that the second derivatives, $m_0''(C)$, $m_1''(C)$ and $d_1''(C)$, are all negative.

3.1.1 Effect of fertility on the probability of divorce

The couple separates at time 1 if the utility when divorced exceeds the marital utility. Let's define q_1 as the probability that the couple will divorce in the second period; then:

$$q_1 = \Pr(U_1^M < U_1^D) = \Pr\{m_1(C) + \theta_1 < d_1(C)\} = \Pr\{\theta_1 < d_1(C) - m_1(C) \equiv \theta^*\} \quad (5)$$

It follows that:

$$q_1 = \int_{-\infty}^{\theta^*} f(\theta_1) d\theta_1 \quad (6)$$

tion; however, since I am interested in studying only the effect of fertility on divorce decision, I do not gain further insights by introducing more variables and complicating the model.

²²I also assume that $m(0) = d(0)$

²³I do not make any assumption on the sign of this relationship; it could be $m_0(C) \leq m_1(C)$.

Differentiating equation (6) with respect to C , I get:

$$\frac{\partial q_1}{\partial C} = f(\theta^*) \left(d'_1(C) - m'_1(C) \right) \quad (7)$$

If the marginal utility of children for the couple when married is greater than when divorced, that is $m'_1(C) > d'_1(C)$, this implies that the higher the number of children C born at time θ , the smaller the *ex-post* probability of divorce ($\frac{\partial q_1}{\partial C} < 0$); the size of the reduction depends upon the marginal utility of children under the two marital statuses, as well as the density function of the match value of the marriage. Thus the presence of children discourages marital disruption.

3.1.2 Effect of the perceived probability of divorce on fertility

However, as mentioned in section 3, causation also runs in the other direction, that is, marital instability exerts effects on fertility.

I want to show that couples with low values of $\hat{\theta}_0$, indicating a strong potential for a low gain to marriage, will tend to give birth to a lower number of children.

In the *first period*, the couple uses the noisy estimate, $\hat{\theta}_0$, to compute an estimate of its divorce probability, \hat{q}_0 , which can be defined as the *ex-ante divorce probability*,²⁴ using the same procedure as in equation (5), I get:

$$\hat{q}_0 = \Pr \left[(U_1^M < U_1^D) / \hat{\theta}_0 \right] = \int_{-\infty}^{\theta^*} f(\theta_1 / \hat{\theta}_0) d\theta_1 \quad (8)$$

After estimating $\hat{\theta}_0$, the couple decides on the number of children to have by maximizing the discounted present value of the utility, given by the utility of being married at time θ plus the expected utility at time 1 :²⁵

²⁴**In other words, it is the probability computed today that a divorce will happen tomorrow.**

²⁵I have assumed a discount rate equal to zero for the sake of simplicity.

$$\max_C U_0^M + E \left(U_1 / \hat{\theta}_0 \right) \quad (9)$$

where $U_1 = \hat{q}_0 U_1^D + (1 - \hat{q}_0) U_1^M$. Since U_1^D is known at time θ , while U_1^M is random because of θ_1 , which is unknown at time θ , equation (9) becomes:

$$\max_C U_0^M + \hat{q}_0 U_1^D + (1 - \hat{q}_0) E \left(U_1^M / \theta_1 > \theta^*, \hat{\theta}_0 \right) \quad (10)$$

where $\theta_1 > \theta^*$ represents the condition to be satisfied for still being married at time 1.

By substituting equation (2), (3) and (4) into (10), the maximization problem for the couple becomes:

$$\max_C m_0(C) + \hat{q}_0 d_1(C) + (1 - \hat{q}_0) [m_1(C) + E(\theta_1 / \theta_1 > \theta^*, \hat{\theta}_0)] \quad (11)$$

The first order condition of the maximization is the choice of C which satisfies:

$$\begin{aligned} & m_0'(C) + \frac{\partial \hat{q}_0}{\partial C} \left[d_1(C) - m_1(C) - E(\theta_1 / \theta_1 > \theta^*, \hat{\theta}_0) \right] \\ + & (1 - \hat{q}_0) \frac{\partial E_0(\theta_1 / \theta^*, \hat{\theta}_0)}{\partial C} + \left[\hat{q}_0 d_1'(C) + (1 - \hat{q}_0) m_1'(C) \right] = 0 \end{aligned} \quad (12)$$

It is possible to show that the third and the fourth terms of equation (12) have equivalent, but opposing effects. The formal proof of their equivalence is provided in Appendix A.

Therefore, equation (12) becomes:

$$m_0'(C) + \hat{q}_0 d_1'(C) + (1 - \hat{q}_0) m_1'(C) = 0 \quad (13)$$

By using the implicit function theorem with equation (13), I get:²⁶

$$\text{sign} \left(\frac{\partial C}{\partial \hat{\theta}_0} \right) = \text{sign} \left(\frac{\partial \hat{q}_0}{\partial \hat{\theta}_0} \left[d_1'(C) - m_1'(C) \right] \right) > 0 \quad (14)$$

²⁶The proof is the following: let's call $F(C, \hat{\theta}_0)$ the left hand side of equation 13. By applying the implicit function theorem to equation 13, I get that the derivative of C with respect to the estimated gain to marriage, $\hat{\theta}_0$ is:

The expected implication is found:²⁷ if $m'_1(C) > d'_1(C)$, as hypothesized in equation (7), couples with more optimistic signals will increase their fertility, while couples with pessimistic signals, indicating a higher *ex-ante* divorce probability, will give birth to a lower number of children (if any).

Based on these implications, the hypothesis that married couples with many children are less likely to separate, and that married couples with *ex-ante* greater separation probabilities are less likely to have many children, can be tested.

4 From the model to the data

Let's specify the functional forms of the expressions $m_0(C)$, $m_1(C)$, and $d_1(C)$ according to the hypotheses made in 3.1. Let's define:

$$m_0(C) = \mu_0 C - \frac{1}{2} C^2 \quad (15)$$

$$m_1(C) = \mu_1 C - \frac{1}{2} C^2 \quad (16)$$

$$d_1(C) = \delta_1 C - \frac{1}{2} C^2 \quad (17)$$

Define $G_1 = U_1^D - U_1^M = d_1(C) - m_1(C) - \theta_1 = (\delta_1 - \mu_1)C - \theta_1$ as the unobservable gain to divorce. Even if G_1 is not observable, the

$$\frac{\partial C}{\partial \hat{\theta}_0} = - \frac{\frac{\partial F(C, \theta_0)}{\partial \theta_0}}{\frac{\partial F(C, \theta_0)}{\partial C_0}}$$

Since what is crucial is the sign of the left-hand side, and since the denominator of the right-hand side has to be negative because of the SOC for a maximum, it follows that the sign of $\frac{\partial C}{\partial \hat{\theta}_0}$ is equal to the sign of $\frac{\partial F(C, \theta_0)}{\partial \theta_0}$.

²⁷The assumption (1) at page 8, according to which higher values of $\hat{\theta}_0$ act to right-shift the couple's subjective density of θ_1 and therefore reduce the expected probability of divorce \hat{q}_0 , is crucial for the identification of the sign of $\frac{\partial \hat{q}_0}{\partial \hat{\theta}_0}$.

couple is observed to divorce ($D_1=1$) or to stay still married ($D_1=0$) in the second period. Let's consider the following model for D_1 :

$$D_1 = \mathbf{1}(G_1 > 0) = \mathbf{1}[\theta_1 < (\delta_1 - \mu_1)C] \quad (18)$$

where $\mathbf{1}$ represents the indicator function.

Therefore, the probability of divorce, given by equation (5), can be expressed as:

$$q_1 = \Pr(D_1 = 1) = \Pr(\theta_1 < (\delta_1 - \mu_1)C) = F_{\theta_1}[(\delta_1 - \mu_1)C] \quad (19)$$

where F represents any function.²⁸

Following the same procedure as equation in 19, the *ex-ante* divorce probability, given by equation (8), becomes:

$$\hat{q}_0 = F_{\theta_1/\hat{\theta}_0}[(\delta_1 - \mu_1)C] \quad (20)$$

and the solution of the maximization problem, given by equation (13), will be:

²⁸ F could be a linear function of the data, and thus we would have a linear probability model, or it could be the cumulative standard normal distribution and we would have a probit model. This last case derives from the following hypotheses on the density functions $f(\hat{\theta}_0)$ and $f(\theta_1/\hat{\theta}_0)$. If I assume that:

- 1) $\hat{\theta}_0 \sim N(0, \tau^2)$
- 2) $\theta_1/\hat{\theta}_0 \sim N(\hat{\theta}_0, \sigma^2)$

it is straightforward to see that $\theta_1 \sim N(0, \sigma^2 + \tau^2)$.

In fact:

- 1) $E(\theta_1) = E(E(\theta_1/\hat{\theta}_0)) = E(\hat{\theta}_0) = 0$;
- 2) $Var(\theta_1) = E(Var(\theta_1/\hat{\theta}_0)) + Var(E(\theta_1/\hat{\theta}_0)) = \sigma^2 + \tau^2$

Hence, equation 19 would become:

$$q_1 = \Phi\left(\frac{\delta_1 - \mu_1}{\sqrt{\sigma^2 + \tau^2}}C\right)$$

where $\Phi(\cdot)$ indicates the cumulative standard normal distribution.

$$C = \frac{1}{2} \left[\mu_0 + \mu_1 + (\delta_1 - \mu_1) \hat{q}_0 \left(C, \hat{\theta}_0 \right) \right] \quad (21)$$

Note that equations 19 and 21 are simultaneous equations.

Most of the literature on divorce has analyzed the effect of fertility on divorce by estimating only equation 19. However, my model clearly shows that such a procedure leads to a biased estimate of this effect because it does not consider the endogeneity of fertility (equation 21). The bias comes from the potential correlation between “number of children” and the error term for the “divorce” relationship, due to two conceptually distinct sources. The first is that families treat the choices of if and how many children to have and the decision to divorce as aspects of a joint decision problem. The second source of correlation is the persistent omitted factors that affect both fertility and marital instability, in which case at least part of the observed relationship between them is spurious. This is the so-called selection bias problem, which implies that those households with children would behave differently from those households with no children, independently of any true causal effect of children on divorce.

Therefore, in order to obtain consistent estimates, I focus on the divorce equation 19 and account for endogenous fertility (equation 21) by using IV methods. In this context, both OLS and probit estimations can be performed, according to the hypotheses made on the function F (see note 28).²⁹

²⁹Alternatively, I could perform maximum likelihood estimation of the structural model of marital dissolution and fertility, by specifying all the unobservable components in the two functions. Very interestingly, by using panel data in this model, it would be possible to infer something on the dynamic process by which individuals learn about the quality of their matches and modify their fertility choices over time.

This represents a promising line of investigations but it goes beyond the scope of this paper and it is left for future research.

5 Estimation strategy

Even if I do not observe the probability of divorce given by equation 19, but only whether the couples divorce or not, however I can estimate a transformation of equation 18. If I assume that F is a linear function of C , and note that $E(D_1) = q_1$,³⁰ with some transformations,³¹ equation 18 can be rewritten as:³²

$$D = \beta C + \theta \quad (22)$$

where θ represents the unobservable gain to divorce and is correlated with C ,³³ and β is the coefficient $(\delta_1 - \mu_1)$ of equation 18. Furthermore, the unobservable gain to divorce, θ , can be decomposed into two parts, as said in note 20: one represents the couple's demographic and social characteristics known at the beginning of marriage, \mathbf{X} , and the second represents some components of the match remained unobserved until later in the marriage, ϵ_1 . Therefore, equation 22 becomes:

$$D = \beta C + \mathbf{X}\gamma + \epsilon \quad (23)$$

In order to estimate equation 23 by IV technique, I need to find a valid instrument, that is a variable that causes some families to have an additional child and others not, but not directly affecting the decision to divorce. From the fertility equation 21, I argue that this variable may be given by one of the observable components of the marginal utility of an additional child at time t , μ_0 , because it appears in equation 21 but not in equation 19 (or 23).³⁴

The component might be the “sex composition of previous children”. This instrument exploits the widely observed phenomenon of

³⁰In fact: $E(D_1) = 1q_1 + 0(1 - q_1) = q_1$

³¹ $D_1 = E(D_1) + (D_1 - E(D_1)) = q_1 + (D_1 - E(D_1)) = (\delta_1 - \mu_1)C + \theta_1 = \beta C + \theta_1$

³²From this moment on, I will leave out the index 1 for D_1 and ϵ_1 to simplify notation, with the remainder that it applies to the second period.

³³More precisely, θ_1 is linked to $\hat{\theta}_0$ that represents a subjective estimate of θ_1 which in turn is correlated with C , as shown in equation 21.

³⁴The possibility of isolating μ_0 in equation (23) relies on the assumption that $m'_0(C) \neq m'_1(C)$.

parental preferences for “balanced” families in terms of the sex composition of their children, at least in developed countries.³⁵ If parents prefer a mixed sibling-sex composition, then having children of the same sex reduces the utility of the existing children, and in turn raises the marginal utility of a new birth, increasing the probability that parents will try to have an additional child.³⁶

As Angrist and Evans (1998) suggest, “because sex mix is virtually randomly assigned, a dummy for whether the sex of the second child matches the sex of the first child provides a plausible instrument for further childbearing among couples with at least two children” (p. 451).³⁷ The instrument can be written as:

$$SameSex = s_1s_2 + (1 - s_1)(1 - s_2) \quad (24)$$

where s_1 and s_2 are dummy variables indicating sex firstborn and second-born children.

However, it should be noted that since *Same Sex* is an interaction term given by the product of the sex of the first two children, it is potentially correlated with the sex of either child. In fact, as Angrist and Evans point out, if we assume that child sex is independent and identically distributed over children, the correlation between *Same Sex* and either s_j is zero only if $E(s_j) = 1/2$.³⁸ If the probability of giving birth to a male child in the sample used is different from $1/2$,³⁹ then there would be some correlation between *Same Sex* and the sex of each child. This correlation represents a problem only if s_j affects divorce for reasons

³⁵This finding is well documented in the demography literature. See Ben-Porath Y. and Welch F. (1976) and Morgan P.S., Lye D.N. and Condran G.A. (1988).

³⁶This idea has been firstly exploited by Angrist and Evans (1998), who study the labor-supply consequences of childbearing explicitly taking into account the endogeneity of fertility.

³⁷However, couples with fewer than 2 children will not be dropped, since they are relevant for the interpretation of the results.

³⁸See Angrist and Evans (1998, p. 460) for the proof.

³⁹Actually, in Table 4 it is shown that the probability of giving birth to a male child is 0,548.

other than family size.⁴⁰ Therefore, it is necessary to include s_1 and s_2 as regressors in the estimating equations, in order to reduce the likelihood of omitted variables bias from these sources.⁴¹

Consequently, the equation to be estimated by IV methods, using *Samesex* as an instrument for C , becomes:

$$D = \beta C + \mathbf{X}\gamma + \eta_1 s_1 + \eta_2 s_2 + \epsilon \quad (25)$$

Now that the problem of identification of the instrument has been solved, the predicted values of the fertility equation can be calculated and then substituted for the original “number of children” variable in the divorce equation 25 (Two-Stage Least-Squares Estimation). In this way, the IV estimates of the effect of the endogenous variable fertility are consistent and causal inferences about the effect of fertility on divorce can be made.

6 Data, methods and descriptive statistics

This section is divided into two parts. In the first part, the data sets used, the variables chosen and the procedure of sample selection are briefly described. In the second part, descriptive statistics and some evidence on the phenomenon of parental preferences for a mixed sibling-sex composition in Germany, the UK and the USA are provided.

⁴⁰As Angrist and Evans suggest, such effects could arise, for example, if the sex of children affects paternal participation in family life, which is higher when all children are boys, generating a lower likelihood of divorce. In general, if parents invest more time in sons than daughters, then in economic terms girls engender less marital-specific capital than boys. Furthermore, effects of sex mix on marital dissolution could also be generated by the fact that boys are more likely than girls to have disabilities (Angrist and Victor Lavy, 1996) and having a disabled child might generate marital distress. Finally, parents may also anticipate fewer long-run benefits from daughters than from sons.

⁴¹If not included as control variables, they would be left in the error terms, violating, in this way, the basic assumption for the application of IV methods, that is, zero correlation between the instrument and the error terms of both fertility and divorce equations.

6.1 The data, the variables and the process of sample selection

The empirical analysis of this section is based on data from Germany, the UK and the USA.⁴² The German data come from the *German Socio-Economic Panel* (GSOEP) in its 95% public-use version.⁴³ The British data come from the *British Household Panel Survey* (BHPS).⁴⁴ The USA data come from the *Panel Study of Income Dynamics* (PSID).⁴⁵

Summary statistics for the dependent variable, covariates and instruments are reported in Tables 2 for the pooled sample.⁴⁶ The dependent variable is the indicator for marital status (divorced in one of the five years of analysis or still married at the end of this period). The covariate of interest is the indicator of the *Number of children* and the instrumental variable for *Number of children* is the indicator *Same Sex*. Since the *Same Sex* instrument can be decomposed into two instruments, also the two indicators *Two boys* and *Two girls* can be used as instruments for *Number of children*.

In order to build the final sample of analysis, I follow this simple procedure of sample selection: first I select only the couples married in

⁴²For the German sample, I only make use of the West German and Foreigners subsamples for the waves 7-12 (1990-1995). This choice is due to the fact that the sample of immigrants has been collected only since 1994, while the East German subsample is excluded because the income variables are not comparable with those of the first two subsamples, at least for the two years after the German reunification in 1989. For the USA sample, I use the five waves 1990-1995. For the British sample, I make use of the first five waves (1991-1996), in order to extract comparable datasets (with the German and the US samples) in terms of the number of years of analysis. However, for Germany and the USA I access a simplified version of their panels, the CNEF Equivalent File 1980-1997, which contains equivalently defined variables for the PSID and for the GSOEP. Since the CNEF Equivalent File 1980-1997 can be merged with the original surveys, I incorporate these constructed variables into my current analyses.

⁴³For more details see Haisken-De New and Frick (1998).

⁴⁴For more details see Rose et al. (1991).

⁴⁵For more details see Martha S.Hill (1992).

⁴⁶Summary statistics for the three countries are available from the author.

1990 (1991 for the UK), then I follow these couples in the next five years and I identify whether they divorced or separated in one of these five years or are still married at the end of the period analyzed. Couples in which one of the two spouses dies are excluded from the sample. Thus, at the end of the fifth year of analysis, the sample is composed of all the couples which were married in 1990 (1991 for the UK sample), and are still married in 1995 (in 1996 for the UK sample) or else divorced or separated in this five-year period.⁴⁷ Each record describes family characteristics, like yearly total household income, and personal characteristics, like age, education, labor earnings, presence of children, and religious affiliation for both partners. After restricting the sample to households with complete records in the critical variables, 7,289 records remain (2,070 for Germany, 1,900 for the UK, and 3,319 for the USA). This constitutes the pooled restricted sample of household observations in the three countries on which the estimation results are based.

6.2 Descriptive statistics

In Table 3, the main household characteristics are reported by partnership outcome. From this table, it can be argued that divorced and married couples are very different with respect to their characteristics; for instance, divorced couples are younger, have higher levels of education, lower household income, the women have lower labor earnings, they have more male second-born children, and more often declare themselves non-Catholics with respect to married couples. All these differences are significant. Moreover, they would seem to declare themselves agnostic, and, more frequently, to have male firstborn children. However, the last two differences are not significant.

Tables 4 and 5 give some insights on the preference for balanced families, which represents the instrument for fertility. They report estimates of the effect of child sex and the sex mix on fertility, similar to

⁴⁷The choice of a period of this length comes from the fact that the decision for divorce or separating usually takes a long time, particularly because of the length of legal procedures.

those in Ben-Porath and Welch (1976) and in Angrist and Evans (1998). They are useful in order to find evidence of whether there is a preference for male first births and whether families with a more equal number of boys and girls are less likely to have another child than those with a more unequal number of boys and girls.

Table 4 shows the fraction of women with at least one child who had a second child, in subgroups categorized by the sex of the first child. It gives an idea about sex preferences in families with one or more children. The third row of this table shows the difference by sex. The data indicate that the fraction of women who had a second child is almost invariant to the sex of the first child. For instance, 62.9% of couples with one girl have a second child with respect to 63.9% of couples with one boy, and this difference is not statistically significant (see third row). Therefore, although “attitudinal surveys suggest many couples would prefer more boys than girls, or prefer their firstborn child to be male” (Angrist and Evans 1998, p. 456), the results in Table 4 suggest that parents are no more or less likely to have a second child if they have a girl first.

Table 5 shows the relationship between the fraction of couples who have a third child and the sex of the first two children. In the first three rows the characteristics of couples in the sample with one boy and one girl, those with two girls and those with two boys are reported. The next two rows show estimates for couples with children of the same sex and for couples with one boy and one girl. The final row reports the differences between the same-sex and mixed sex group averages. From Table 5, I observe that couples with two children of the same sex are more likely to have a third child than couples who have one boy and one girl: only 31.7% of couples with one boy and one girl have a third child compared to 37.5% for couples with two girls or two boys. This difference is statistically significant (see row 6).

7 Estimation results

In this section, I first present OLS estimations, using two different measures of fertility; second, I present the Wald estimates, which give an idea of how the instruments identify the effect of children on marital dissolution. Third, I provide some evidence on the quality of the instrument chosen, *Same Sex*, and report first-stage results linking sex mix and fertility. Then, I report and interpret the estimation results of the effect of fertility on divorce from two-stage least squares instrumental variables regression (2SLS), which is considered as a more sophisticated statistical technique than Wald estimation; finally, I compare these results with those obtained from OLS estimations.

7.1 OLS estimates⁴⁸

In the OLS-estimates two measures on fertility are considered: the first one considers children between 0 and 18 years old, while the second measure disentangles the fertility information by considering young children between 0 and 12 years old and adolescents between 13 and 18 years old (see Table 6). In the first case the coefficient on the number of children 0-18 years old is -0.004 (s.e. 0.004), while in the second case the coefficient on the number of children 0-12 years old is -0.008 (s.e. 0.004) and the one on children 13-18 is 0.005 (s.e. 0.006). It is evident from these regressions that the negative relationship between fertility and divorce comes entirely from the presence of young children between 0 and 12 years old in the household. In particular, this result indicates that the presence of an additional child reduces the probability of divorce by 0.8 percentage points. It confirms the estimates previously reported in the literature on fertility and divorce, according to which younger children

⁴⁸Since the dependent variable is binary, a probit or logit specification of the model would have been more correct, but I have preferred to report the OLS estimates for comparability with the 2-SLS results. However, I have computed probit and logit estimates for robustness check, and I have found no significant differences with the results reported in Table 6. The results are available from the author.

discourage marital dissolution more than older children. At the light of this result, in the rest of the analysis I will focus on the effect of the presence of young children on marital dissolution.

However, as already pointed out in the previous section, OLS estimates can generally not be considered as estimates of causal effects because of the problem of self-selection and the correlation of fertility with unobservable characteristics such as perception of stability of marriage, love, etc., that make these estimates biased. Therefore, I turn to IV techniques which provide unbiased estimates of the causal effect of fertility on divorce.

7.2 Wald estimates

Because sibling-sex composition is virtually randomly assigned, I can use the Wald estimates to illustrate how the instrument identifies the effect of fertility on marital dissolution.⁴⁹ The starting point is the simple bivariate regression model given by equation 25 without covariates:⁵⁰

$$D_i = \beta C_i + \epsilon_i \quad (26)$$

It can be estimated by using Wald estimation techniques; in fact, using the binary instrument, *Same Sex*, the IV estimate of β in equation 26 can be written as:

$$\beta_{IV} = (\overline{D}_1 - \overline{D}_0) / (\overline{C}_1 - \overline{C}_0) \quad (27)$$

where \overline{D}_1 is the mean of D_i for those observations with *SameSex* = 1; the same definition applies to the other terms. The numerator and the denominator of equation 27 are the reduced-form relationships respectively between the dichotomous variable identifying marital dissolution D_i and the instrument *SameSex* and between fertility measure C_i and *SameSex*. The first row of Table 7 shows the denominator of the Wald

⁴⁹In this paragraph I follow Angrist and Evans (1998, section II A) who use the same strategy to identify the effect of fertility on parents' labor supply.

⁵⁰The index i refers to a generic couple.

estimate, $\bar{C}_1 - \bar{C}_0$; it indicates that the effect of the *SameSex* instrument on *Number of children 0 – 12* is 0.993. It means that couples with children of the same sex are more likely to have an additional child than couples with one boy and one girl. The second row reports $\bar{D}_1 - \bar{D}_0$ using the *SameSex* instrument as regressor. It shows that couples with two children of the same sex have a higher probability of divorce than those with mixed-sex siblings. The Wald estimate in the second column, calculated by dividing $\bar{D}_1 - \bar{D}_0$ by $\bar{C}_1 - \bar{C}_0$, shows that having an additional child increases the probability of divorce by 2.04 percentage points.

However, even if the Wald estimates are very useful to sketch how the sex-mix IV strategy identifies the effect of fertility on the probability of divorce, in the rest of the paper I focus on two-stage least-squares (2SLS) estimates of regression models relating divorce to fertility and other exogenous covariates. This choice is due to three reasons (suggested by Angrist and Evans, 1998). First, even if the instrument is not correlated with the exogenous covariates, controlling for them can give more precise estimates if the fertility variable is approximately constant across groups. Second, 2SLS can be used to control for any additive effects of child sex when using *Same Sex* as an instrument.⁵¹ Third, by 2SLS, the *Same Sex* instrument can be decomposed into two instruments, *Two boys* [s_1s_2], and *Two girls* [$(1 - s_1)(1 - s_2)$], generating an overidentified model. This decomposition allows me to investigate further the hypothesis that the divorce consequences of childbearing depend on whether *Same Sex* equals *Two boys* or *Two girls*.

7.3 First-stage results

Table 8 gives an indication of how well the instrument *Same Sex* explains fertility. In particular, I examine how the sex of previous children

⁵¹As already explained in section 5, by adding two dummies for the sex of each child as regressors the likelihood of omitted-variables bias from these two sources can be reduced. However, as Angrist and Evans (1998) notice, “controls for additive effects can only eliminate bias from omitted variables with effects that are additive in the number of children” (p. 461).

exogenously alters fertility. The estimates are for linear equations which include indicators of the sex of previous children, mean age, mean education, household total income (in log), wife labor earnings (in log), Catholic affiliation and no religion for both persons in the couple, and living in Germany or in the USA (the omitted category is living in the UK). The results reveal that having children of the same sex has a significant and positive effect on the probability of having an additional child both without (0.98) and with covariates (0.55)(columns 1-2). Note that when the effects of sex mix are allowed to differ by sex (columns 3), the effect of *Same Sex* is much larger for girls than for boys (1.03 respect to 0.43).⁵²

Furthermore, the instrument *Same Sex* is clearly correlated with fertility (0.400), but only slightly (and positively) correlated with the outcome “divorce” (0.031).⁵³

7.4 2-SLS preliminary results

In table 9 the two sets of 2SLS estimates using respectively *Same Sex* and the dummies *Two boys* and *Two girls* as instruments are presented.

First-stage F-statistics and partial R^2 measures are also reported as a diagnostic tool following the suggestions of Bound et al. (1995), and they indicate the high quality of the instrument(s).

By comparing the first column of table 9 (2SLS estimates) with the second column of table 6 (OLS estimates), I see that the use of sibling sex composition as an exogenous determinant of fertility yields IV estimates of the effect of fertility on marital instability to be markedly different from the estimates obtained under strict exogeneity, implying a positive effect of young children on divorce (0.006 with respect to -0.008 estimated

⁵²Note that in this last case either s_1 or s_2 must be dropped from the list of covariates because $s_1, s_2, s_1s_2, (1 - s_1)(1 - s_2)$ are linearly dependent.

⁵³This result is important because the instrument should not be correlated with the outcome in order to generate unbiased estimates. However, this result can not be considered as a test of the *exclusion restriction* assumption.

by a conventional ordinary least squares procedure).⁵⁴ The same result holds if *Two boys* and *Two girls* are used as instruments (0.007).

Since the dependent variable (marital status) is binary, I have also performed a 2SLS with a probit second stage and opportunely corrected standard errors. In this case, the result does not change significantly (the coefficient is 0.016 with a standard error of 0.053).⁵⁵

However, the standard errors of the IV estimates are relatively large; therefore, this coefficient is not significantly different from zero. A zero coefficient implies no causal effect of fertility on divorce after controlling for the potential endogeneity of fertility. In terms of the model presented in section 3.1.1, this means that, once the effect of the couple's perceived probability of dissolution on fertility is controlled for, an effect of the number of young children on the *ex-post* probability of divorce is no longer expected.

The *Hausman test* is then used to test whether the difference between the coefficient of children of the instrumental variable regression and standard OLS is significant. The result of the *Hausman test* suggests that there is not enough evidence to refuse the hypothesis H_0 of no differences between the two coefficients. However, even if this difference is not statistically significant, it is in size important from an economic point of view. In fact, the results show that the coefficient of *Number of children 0-12* goes from a negative OLS estimate (-0.008) to a positive IV estimates (0.006, using *Same Sex* as instrument). Given that the rate of divorce in the pooled sample is 0.12 (see Table 2), the mentioned difference is 17% of this value, that is to say, going from the OLS estimate to

⁵⁴Looking at other coefficients in the two tables, it can be seen that they are of similar signs and significance: a higher average age of the couple, a higher average education of the couple, being Catholic, a higher household income, and having a male firstborn reduce the probability of divorce, while a higher wife labor earning, being agnostic and having a male second-born seem to increase the probability of divorce. Furthermore, living in Germany reduces the probability of divorce and living in the USA increases the probability of divorce with respect to the reference category "living in the UK". It should be noted, however, that some of these coefficients are not significant.

⁵⁵Tables and programs of the probit-2SLS are available from the author.

the IV estimate, the probability of divorce increases by 17% with respect to the rate of divorce in the sample.

Furthermore, as it can be seen in Table 8, couples with two girls are more likely than couples with two boys to have another child. Hence, in the first-stage relationships I find different results according to the choice of the instrument. However, the 2SLS estimates in column (2) of Table 9 show that no additional insights are gained by separating the two components of *Same Sex* because they neither change the coefficient estimates nor increase their efficiencies. In the last row of Table 9, the result of the overidentification test associated with the use of *Two boys* and *Two girls* as instruments is reported:⁵⁶ the p -value suggests that there are no differences in using one instrument or the other.

Although imprecise, the instrumental variables results presented in Table 9 make it clear that a conventional OLS estimation strategy yields a downward-biased estimate of the "true" effect of fertility on divorce, the former negative and the latter positive. This finding seems directly at odds with many papers which state that couples with young children tend to divorce less.

The imprecision of the IV estimates suggests that alternative estimation method has to be explored.

I also evaluate the robustness of the results. In particular, I evaluate the sensitivity of the results to different fertility measure choices, to the choice of different subsamples of the original data set, and I investigate whether the results change when I add controls.⁵⁷ The last robustness check is particular important because it allows to investigate the problem of the potential endogeneity of the female education and earning.⁵⁸ I

⁵⁶It jointly tests for a difference between 2SLS estimates using only *Two boys* and 2SLS estimates using only *Two girls*, as suggested by Angrist and Evans (1998).

⁵⁷This type of robustness test is proposed by Bertrand, Luttmer, and Mullainathan [1998, p.26] who suggest that, "if unobservable characteristics about individuals drove our results, one would expect that increasing the set of unobservables characteristics by treating observable characteristics as unobservable would have a large impact on the estimate" of fertility effects.

⁵⁸I thank Prof. Daniela Del Boca for making me notice it.

begin with a simple regression which has only the fertility measure and the mean age of the couple (row (1) of table 10). The coefficient in row (1) for the OLS estimates is lower than the corresponding coefficient of the original regression in table 6, while the coefficients from 2SLS estimates do not change respect to the coefficients of the original regressions in table 9). Then I add controls that are clearly exogenous: mean age squared, being German, being American, being Catholic, being agnostic, having a male firstborn child, having a male second-born child and the household total income (in logarithms). The effect of fertility on divorce (row (2)) is even stronger once I introduce these controls. In row (3), I add average years of education of the couple and the labor earnings of the wife (in log). The choice of isolating these last two covariates from the others depends on the fact that they are potentially endogenous because they might be partly determined by the perception of a high probability of divorce.⁵⁹ However, the inclusion of the education and wife’s earnings controls does not change the coefficient (row (3)). The estimation results seem quite robust to all the checks above mentioned and especially to the potential endogeneity of female education and earnings.⁶⁰

8 Conclusions and further research

In this paper, a model which relates a couple’s divorce probability and their fertility decision is developed and then used to estimate the relationship between them. Any credible analysis of the causal link between fertility and divorce requires an exogenous source of variation in fertility choices. IV estimation methods are used in order to take into account the potential endogeneity of fertility and the sex of previous two children is explored as an exogenous determinant of fertility.

The five waves 1990-1995 from the German Socio-Economic Panel (GSOEP) and the Panel Study of Income Dynamics (PSID), and the five

⁵⁹Women who perceive a high probability of separation at some time in the future, have greater incentives to make investments in schooling or in career and to postpone having children.

⁶⁰All the results of the tests are available from the author.

waves 1991-1996 from the British Household Panel Survey (BHPS) allow me to test the model's predictions.

Two important conclusions emerge from the analysis. First, the standard approach which does not instrument fertility (OLS estimates) leads to a negative and significant estimate of the impact of exogenous changes of fertility on marital dissolution. Second, IV estimates that exploit the fertility consequences of sibling sex composition on marital instability contradict the OLS results. For instance, the implied instrumental variables estimate of the effect of fertility on divorce (0.006) is substantially above and of different sign with respect to the probability of divorce estimate by a conventional ordinary least squares procedure (-0.008). Nevertheless, the standard errors of the IV estimates are relatively large and one cannot reject the hypothesis that the difference between the IV and OLS estimates is not statistically significant. However, this difference is economically meaningful, thus imposing the exploitation of alternative estimation method.

As already mentioned in the text, further research into the topic can be undertaken both theoretically and empirically.

From a theoretical point of view, the divorce decision can be analyzed in the context of a bargaining model where two agents who optimize separate utility functions, who have differing utility levels outside marriage, and who interact in fertility and divorce decisions, are specified.

Empirically, as said before, alternative estimation methods have to be identified. In particular "single equation" estimation techniques by using propensity scores methods can be explored (see Rosenbaum P.R. and Rubin D.B.,1983, and Imbens G., 1999);

The comparison of results from these different approaches will be the focus of future research.

9 Appendix A

Proof. In this section a proof of the equivalence of the third and the fourth term of equation (12) is provided.

To begin, note that:

$$E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0) = \frac{\int_{\theta^*}^{\infty} \theta_1 f(\theta_1/\hat{\theta}_0) d\theta_1}{\int_{\theta^*}^{\infty} f(\theta_1/\hat{\theta}_0) d\theta_1} = \frac{-\int_{-\infty}^{\theta^*} \theta_1 f(\theta_1/\hat{\theta}_0) d\theta_1}{1 - \hat{q}_0} \quad (28)$$

Therefore:

$$\begin{aligned} \frac{\partial E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)}{\partial C} &= \frac{\partial}{\partial C} \frac{-\int_{-\infty}^{\theta^*} \theta_1 f(\theta_1/\hat{\theta}_0) d\theta_1}{1 - \hat{q}_0} = \\ &= \frac{-\theta^* f(\theta^*/\hat{\theta}_0) (1 - \hat{q}_0) \frac{\partial \theta^*}{\partial C} - \frac{\partial \hat{q}_0}{\partial C} \int_{\theta^*}^{\infty} \theta_1 f(\theta_1/\hat{\theta}_0) d\theta_1}{(1 - \hat{q}_0)^2} \end{aligned} \quad (29)$$

From equation (8) it follows that:

$$\frac{\partial \hat{q}_0}{\partial C} = f(\theta^* | \hat{\theta}_0) \frac{\partial \theta^*}{\partial C} \quad (30)$$

Substituting equation (30) in (29) and with some simplifications we get:

$$\begin{aligned} \frac{\partial E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)}{\partial C} &= \frac{-f(\theta^* | \hat{\theta}_0) [\theta^* - E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)]}{(1 - \hat{q}_0)} \frac{\partial \theta^*}{\partial C} = \\ &= -\frac{\frac{\partial \hat{q}_0}{\partial C} \theta^* - E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)}{(1 - \hat{q}_0)} \end{aligned} \quad (31)$$

Substituting equation (31) into equation (12), we get:

$$m'_0(C) + \frac{\partial \hat{q}_0}{\partial C} [d_1(C) - m_1(C) - E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)] \\ - (1 - \hat{q}_0) \frac{\partial \hat{q}_0}{\partial C} \frac{\theta^* - E(\theta_1 | \theta_1 > \theta^*, \hat{\theta}_0)}{(1 - \hat{q}_0)} + [\hat{q}_0 d'_1(C) + (1 - \hat{q}_0) m'_1(C)] = 0$$

Finally, since $\theta^* = d_1(C) - m_1(C)$, it is easy to see that the third and the fourth term of equation (32) cancel out. ■

References

- [1] Amemiya T., (1974), “*Multiple regression and simultaneous equation models when the dependent variables are truncated normal*”, *Econometrica* 42: 999-1012.
- [2] Andersson G., (1997), “*The Impact of Children on Divorce Risks of Swedish Women*”, *European Journal of Population* 13: 109-145.
- [3] Angrist J.D. and Evans W., (1998), “*Children and their parents’ labor supply: evidence from exogenous variation in family size*”, *The American Economic Review* 88(3): 450-477.
- [4] Angrist J.D. and Imbens G.W., (1995), “*Two-stage least squares estimation of average causal effects in models with variable treatment intensity*”, *Journal of the American Statistical Association*, 90 (430), 431-442
- [5] Angrist J.D., Imbens G.W and Rubin D.W., (1996), “*Identification of causal effects using instrumental variables*”, *Journal of the American Statistical Association*, 91 (434), 444-472
- [6] Angrist J.D. and Lavy V., (1996), “*The effect of ten childbearing and single parenthood on childhood disabilities and progress in school*”, NBER Working Paper n°5807.
- [7] Becker G.S., (1998); *L’approccio economico al comportamento umano*; il Mulino

- [8] Becker G.S., Landes E.M. and Michael R.T, (1977), "*An economic analysis of marital instability*", Journal of Political Economy 85(1): 1141-1187.
- [9] Becker G.S., (1991); A Treatise on the Family; Harvard University Press.
- [10] Becker G.S., (1974), "*A theory of marriage: Part II*", Journal of Political Economy 82(2): S11-S26.
- [11] Ben-Porath Y. and Welch F., (1976), "*Do sex preferences really matter?*", Quarterly Journal of Economics 90 (2): 285-307.
- [12] Bertrand M., Luttmer E.F.P. and Mullainathan S., (1998), "*Network effects and welfare cultures*", NBER Working Paper n°6832.
- [13] Boheim R., (1999), "*Unemployment, financial surprises and separation*", Institute for Social and Economic Research, University of Essex.
- [14] Brien M.J., Lillard L.A. and Stern S., (1999), "*Cohabitation, Marriage, and Divorce in a Model of Match Quality*", mimeo, University of Virginia.
- [15] Bound J., Jaeger D.A. and Baker R.M., (1995), "*Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak*", Journal of the American Statistical Association, 90 (430), 443-450
- [16] Brown M., (1999), "*Divorce Rates and Investment in Marriage-Specific Assets*", New York University.
- [17] Browning M., (1992), "*Children and household economic behavior*", Journal of Economic Literature.
- [18] Cherlin A., (1977), "*The effect of children on marital dissolution*", Demography 14: 265-72.

- [19] Cerrasco R., (1998), “*Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labour Participation*”, Insitute of Economics, University of Copenhagen, Working Paper No. 9805.
- [20] DIW (ed.), (1998), The German Socio-Economic Panel (GSOEP) Desktop Companion, Berlin.
- [21] EUROSTAT, (1997), Demographic Statistics.
- [22] Green W.H., (1987), *Econometric Analysis*, 3rd. edition, Prentice Hall.
- [23] Gruber J., (2000), ”*Is making divorce easier bad for children? The long run implications for unilateral divorce*”, NBER Working Paper n°7968.
- [24] Hamermesh D.S.,(1999); “*The Art of Labormetrics*”, NBER Working Paper 6927.
- [25] Hill Martha S., (1992), *The panel study of income dynamics: a user’s guide*, Harvard University
- [26] Holland P.W., (1986), “*Statistics and causal inference*”, *Journal of the American Statistical Association*, 81 945-960.
- [27] Koo H and Janowitz B., (1983), “*Interrelationship between fertility and marital dissolution: results of a simultaneous logit model*”, *Demography* 20 (2): 129-145.
- [28] Jensen P. and Smith N., (1990), “*Unemployment and marital dissolution*”, *Journal of Population Economics* 3: 215-229.
- [29] Lillard L.A. and Waite L.J. (1993), “*A Joint Model of Marital Child-bearing and Marital Disruption*”, *Demography* 30: 653-681.
- [30] Imbens G., (1999), “*The Role of Propensity Score in Estimating Dose Response Functions*”, NBER Technical Working Paper 237.

- [31] Murphy R.D. (1999), " *A good man is hard to find*", Working Paper E98-02 Virginia Polytechnic Institute and State University.
- [32] McLanahan S. and Sandefur G., (1994), *Growing up with a single parents*, Cambridge, MA: Harvard University Press.
- [33] Morgan S. and Rindfuss R., (1985), " *Marital Disruption: Structural and Temporal Dimensions*", *American Journal of Sociology* 90: 1055-1077.
- [34] Morgan P.S., Lye D.N. and Condran G.A. (1988), " *Sons, Daughters and the Risk of Marital Disruption*", *American Journal of Sociology* 94, n°1: 110-129.
- [35] Nakamura A. and Nakamura M., (1992), " *The econometrics of female labor supply and childrens*", *Econometrics Reviews* 11(4): 1-71.
- [36] Ono H., (1998), " *Husbands' and Wives' Resources and Marital Dissolution*", *Journal of Marriage and the Family* 60: 674-689.
- [37] Peters E., (1986), " *Marriage and Divorce: Informational Constraints and Private Contracting*", *The American Economic Review* 76: 437-454.
- [38] Peters E., (1993), " *The Importance of Financial Considerations in Divorce Decisions*", *Economic Inquiry* XXXI: 71-86.
- [39] Picot G., Zyblock M., Pyper W., (1999) " *Why do children move into and out of low income: changing labour market conditions or marriage and divorce?*", *Anaytical Studies Branch-Research Paper Series*, Statistics Canada n°132.
- [40] Rose D. and 14 others, (1991), " *Micro-social change in Britain: an outline of the role and objectives of the British Household Panel Survey*", *Working papers of the ESRC Research Centre on Micro-social Change*, Paper I, Colchester, University of Essex.

- [41] Rosenbaum P.R. and Rubin D.B., (1983), "*The central role of the propensity score in observational studies for causal effects*", *Biometrika* 70: 41-55.
- [42] Sullivan T., (1995), "*Ex ante divorce probability and investment in marital-specific assets: an application to home ownership*", University of Maryland, Working Paper Series No. 95-5.
- [43] Thornton A, (1977), "*Children and marital stability*", *Journal of Marriage and the Family* 39: 531-540.
- [44] Waite L., Booth A. and Edwards J., (1986), "*Children and marital happiness: why the negative relationship?*", *Journal of Family Issues* 7: 131-148:
- [45] Waite L. and Lillard L., (1991), "*Children and marital disruption*", *American Journal of Sociology* 96: 930-53.
- [46] Waite L., Haggstrom G. and Kanouse D., (1985), "*Parenthood and marital stability*", *American Sociological Review* 50: 850-57.
- [47] Weiss, Y and Willis, J.R., (1997), "*Match Quality, New Information and Marital Dissolution*", *Journal of Labor Economics* 15: S293-S329.
- [48] Weiss, Y. (1997), "*The formation and dissolution of families: why marry? Who marries whom? And what happens upon marriage and divorce*" in Stark, O. and Rosenzweig R. (eds.) *Handbook of Population Economics*: 82-123.
- [49] Willis R, (1994), "*A theory of out-of-wedlock childbearing*", unpublished manuscript.

Table 1: Crude divorce rate (per 1000 population)

	B	D	E	F	I	NL	P	UK	USA
1960	0.5	1.0	-	0.7	-	0.5	0.1	0.5	2.2
1965	0.6	1.1	-	0.7	-	0.5	0.1	0.7	2.5
1970	0.7	1.3	-	0.8	-	0.8	0.1	1.1	3.5
1975	1.1	1.9	-	1.1	0.2	1.5	0.2	2.3	4.8
1980	1.5	1.8	-	1.5	0.2	1.8	0.6	2.8	5.2
1985	1.9	2.3	0.5	1.9	0.3	2.3	0.9	3.1	5
1990	2.0	2.0	0.6	1.9	0.5	1.9	0.9	2.9	4.7
1991	2.1	1.7	0.7	1.9	0.5	1.9	1.1	3.0	4.7
1992	2.2	1.7	0.7	1.9	0.5	2.0	1.3	3.0	4.7
1993	2.1	1.9	0.7	1.9	0.4	2.0	1.2	3.1	4.8
1994	2.2	2.0	0.8	2.0	0.5	2.4	1.4	3.0	4.6
1995	3.5	2.1	0.8	2.0	0.5	2.2	1.2	2.9	4.4
1996	2.8	2.1	0.8	2.1	0.6	2.3	1.4	:	:
1997	2.6	:	:	:		2.2	1.4	:	:

Source: EUROSTAT- Demographic Statistics 1997; *For the USA the source is the

U.S. Bureau of the Census International Data Base

Notes: D includes in all years data on the former GDR

Table 2: Descriptive statistics for the pooled sample (sample size 7289)

Variable	mean	st.dev	min	max
DIVORCE	0.12	0.32	0	1
FERTILITY				
#children	1.26	1.18	0	9
children aged 0 to 12	0.75	1.01	0	7
children aged 0 to 18	1.04	1.16	0	9
children aged 13 to 18	0.29	0.60	0	4
Same Sex	0.21	0.41	0	1
Boyfirst	0.36	0.48	0	1
Boysecond	0.21	0.41	0	1
Two boys	0.12	0.32	0	1
Two girls	0.09	0.29	0	1
CONTROL VARIABLES				
Husband's age	44	13.6	18	91
Wife's age	41.4	13.3	16	85
Mean age	42.7	13.3	17.5	85.5
Husband's education	12	2.6	1	19.5
Wife's education	11.7	2.3	2	19.5
Mean Education	11.9	2.2	2	19
Family gross income	69723	81743	0	1949940
Husband's wage	43085	59444	0	1829673
Wife's wage	21771	44750	0	923670
Roman Catholic	0.18	0.38	0	1
No religion	0.07	0.25	0	1
Germany	0.28	0.45	0	1
UK	0.27	0.44	0	1
USA	0.46	0.5	0	1

Notes: : Divorce: = 1 if the couple divorces or separates in one of the 5 years, 0 if still married after 5 years; Same Sex: = 1 if the first two children are of the same sex, 0 otherwise; boy1st: = 1 if the first child is male, 0 otherwise; boy2nd: = 1 if the second child is male, 0 otherwise; twoboys: = 1 if the first two children are boys, 0 otherwise; twogirls: = 1 if the first two children are girls, 0 otherwise; catholic: = 1 if husband and wife are both Roman Catholic, 0 otherwise; no Religion: = 1 if husband and wife have no religion, 0 otherwise; monetary variables are in EURO.

Table 3: Household characteristics by partnership outcome)

Variables	DIV=0(mean)	DIV=1(mean)	t-test(t-statistics)
Mean age	43.65	35.17	17.6
Mean Education	11.84	12.04	-2.55
Household income	71793	53470	6.07
Wife's labor earnings	22201	18390	2.30
Boy first	0.36	0.37	-0.71
Boy second	0.20	0.25	-2.99
No religion	0.065	0.079	-1.07
Roman catholic	0.18	0.09	6.52
Total observations	6466	823	

The full sample is composed by 7289 couples

Table 4: Fraction of couples with one child who had another child, by sex of first child)

<i>Sex of first child in families with one or more children</i>	<i>Fraction of sample</i>	<i>Fraction that had another child</i>
(1) one girl	0.452 (0.070)	0.629 (0.010)
(2) one boy	0.548 (0.070)	0.639 (0.009)
Difference (1)-(2)	-	-0.009 (0.014)

Notes: The sample size is 4973 and it includes couples married in the first year of analysis who have at least one child. Standard errors are reported in parentheses.

Table 5: Fraction of couples with twochildren who had another child, by sex of first two children)

<i>Sex of first two children in families with two or more children</i>	<i>Fraction of sample</i>	<i>Fraction that had another child</i>
one boy, one girl	0.494 (0.010)	0.317 (0.012)
Two girls	0.224 (0.007)	0.400 (0.018)
Two boys	0.281 (0.008)	0.355 (0.016)
(1) one boy, one girl	0.494 (0.010)	0.317 (0.012)
(2) both same sex	0.506 (0.009)	0.375 (0.012)
Difference (1)-(2)	-	-0.057 (0.017)

Notes: The sample size is 3156 and it includes couples married in the first year of analysis who have at least two children. Standard errors are reported in parentheses.

Table 6: OLS estimates of divorce equation)

	(1)	(2)
# children 0-18 years	-0.004 (0.004)	-
# children 0-12 years	-	-0.008~ (0.004)
# children 13-18 years	-	0.005 (0.006)
other covariates	yes	yes

Notes: The sample is composed by 7289 couples. Other covariates in the model are indicator for male firstborn and male secondborn, mean of age, mean of education, household income, wife labor earnings, spouses agnostic, spouses Catholic, being German and being American, being British is the omitted category. Robust standard errors in parentheses

Table 7: Wald estimates of divorce model

Variable	Mean Difference by <i>Same Sex</i>	Wald Estimates
Number of Children 0-12	0.993 (0.026)	-
Probability of Divorce	0.0202 (0.009)	0.0204 (0.009)

Notes: The sample is composed of 7289 couples. Robust standard errors in parentheses

Table 8: OLS Estimates of Number of Young Children - Fertility Equation)

Independent variable	Children between 0-12 years old					
	Model 1		Model 2		Model 3	
	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
Same Sex	0.983	(0.026)	0.554	(0.025)	-	-
Two boys	-	-	-	-	0.430	(0.032)
Two girls	-	-	-	-	1.033	(0.033)
Boy 1 st	-	-	0.231	(0.020)	0.420	(0.023)
Boy 2 nd			0.445	(0.025)	-	
Mean Age			-0.035	(0.005)	-0.033	(0.005)
Age ²			0.000	(0.000)	0.000	(0.005)
Mean Education			0.022	(0.005)	0.022	(0.005)
Household income (lg)			-0.048	(0.008)	-0.046	(0.009)
Wife L. Earnings (lg)			-0.042	(0.002)	-0.038	(0.002)
No Religion			-0.067	(0.037)	-0.088	(0.038)
Roman Catholic			-0.020	(0.024)	0.010	(0.025)
Germany	-0.030	(0.028)	-0.072	(0.026)	-0.075	(0.027)
Usa	0.225	(0.025)	0.126	(0.025)	0.135	(0.025)
Constant	0.446	(0.021)	2.411	(0.121)	2.350	(0.122)
R ²	0.174		0.4150		0.403	

Notes: The sample is composed of 7289 couples. Robust standard errors in parentheses with $p < 0.01 = **$, with $p < 0.05 = *$ and $p < 0.1 = \sim$. For variable definition see Table 2.

Table 9: 2SLS Estimates of Divorce Equation1)

Estimation Method	2SLS		2SLS	
	coefficient	s.e.	coefficient	s.e.
Instrument for <i>Number of children between 0-12 years</i>	<i>Same Sex</i>		<i>Two boys</i>	
			<i>Two girls</i>	
Independent variable				
<i>Number of children between 0-12</i>	0.006	(0.018)	0.007	(0.011)
<i>Boy 1st</i>	-0.008	(0.009)	-0.007	(0.009)
<i>Boy 2nd</i>	0.009	(0.015)	-	
<i>Mean Age</i>	-0.013	(0.002)	-0.013	(0.002)
<i>Age²</i>	0.000	(0.000)	0.000	(0.000)
<i>Mean Education</i>	0.009	(0.002)	0.009	(0.002)
<i>Household income (log)</i>	-0.013	(0.004)	-0.013	(0.004)
<i>Wife Labor Earnings (log)</i>	0.003	(0.001)	0.003	(0.001)
<i>No Religion</i>	0.011	(0.015)	0.011	(0.015)
<i>Roman Catholic</i>	-0.039	(0.010)	-0.039	(0.010)
<i>Germany</i>	-0.016	(0.010)	-0.016	(0.011)
<i>Usa</i>	0.081	(0.010)	0.081	(0.010)
<i>Constant</i>	0.687	(0.064)	0.684	(0.054)
1 st stage F	430.09		410.17	
Partial R ²	0.415		0.404	
Overidentification test (<i>p</i> -value)			0.9996	

Notes: The sample is composed of 7289 couples. Standard errors in parentheses with $p < 0.01 = **$, with $p < 0.05 = *$ and $p < 0.1 = \sim$. For variable definition see Table 2.

Table 10: Sensitivity of Results to Addition of Controls)

Dependent Variable: Divorce

Reported: Coefficient of Number of Children between 0 and 12 years

Estimation Method	OLS	2SLS	2SLS
Instrument for <i>Number of children between 0-12 years</i>	-	Same Sex	Two boys, Two girls
Controls:			
(1) fertility measure + mean age of the couple	-0.005 (0.004)	0.006 (0.011)	0.006 (0.010)
(2) (1) + exogenous controls	-0.009 (0.004)	0.006 (0.013)	0.003 (0.011)
(3) (2) + education + labor earnings of the wife	-0.008 (0.004)	0.006 (0.018)	0.007 (0.011)

Notes: Standard errors in parentheses with $p < 0.01 = **$, with $p < 0.05 = *$ and $p < 0.1 = \sim$.

. For variable definition see Table 2