EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

EUI Working Paper **ECO** No. 2003/4

Propensity Score Estimates of the Effects of Fertility on Marital Dissolution

Daniela Vuri

BADIA FIESOLANA, SAN DOMENICO (FI)

Published in Italy in March 2003
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

Propensity Score Estimates of the Effect of Fertility on Marital Dissolution

Daniela Vuri (YE)*†
European University Institute
IZA

January 16, 2002

Abstract

In recent years many studies have reported significant empirical associations between fertility and marital dissolution. Whether this is a causal effect or only a correlation is not clear. This issue is explored by using matching methods. First the effect of "having children" (binary treatment) on marital disruption is investigated. Then, the method is extended to the case of "number of children in the household" (multi-valued treatment). The main findings indicate that parents do not divorce less in the presence of children but they only postpone the decision to divorce until children get older.

Keywords: Fertility; Marital dissolution; Propensity score methods; Counterfactual
Jel Classification: C12, C2, J12, J13

1 Introduction

In recent years many studies have reported significant empirical associations between fertility and marital dissolution. Whether this is a causal effect or only a correlation is not clear. The goal of this empirical analysis is to find out whether there is a true causal effect of fertility on marital dissolution. If fertility were randomly assigned to the population of married couples, then the observed marital dissolution differential by fertility could be interpreted as a causal effect. However, as shown in Vuri (2001), the marriage continuation probability increases as the number of children increases, but at the same time the *potential* stability of the marriage may

*Department of Economics, European University Institute, Via dei Roccettini 9, I-50016 San Domenico di Fiesole (FI), Italy. E-mail: vuri@iue.it. Fax: +39-055-4685202. Tel: +39-328-8524310

†I would like to thank Andrea Ichino for comments and guidance. Many thanks to participants at the Spring Meeting of Young Economists 2001 in Copenhagen, at the ESPE 2001 Conference in Athens, at the BHPS 2001 Conference in Colchester, at the EEA Conference 2001, and to seminar participants at the EUI, especially Anna Sanz de Galdeano, with whom I programmed the propensity scores methods, and Sascha Becker. A special thank to Erich Battistin who gave me the initial hints to program the propensity score methods and very useful comments. All errors are mine.

affect the arrival of children. Therefore, fertility may not be exogenous to the decision of marital dissolution.

There is a substantial body of literature studying the effect of fertility on marital dissolution and it can be divided into two categories based on the methodological approach used: studies considering fertility as an exogenous variable, and studies addressing the problem of endogeneity of fertility.

Studies in the first category show that children increase the stability of their parents' marriage throughout their preschool years, while children born before marriage increase significantly the chances that the couple will dissolve (Becker, Landes and Michael 1977, White and Lillard 1991, Peter 1986, Ono 1998). The positive effect of having children on marital stability does not seem to hold in case of high number of children (Thornton, 1977).

However, this literature is not satisfactory because it neglects the potential problem of endogeneity of fertility. This implies that if fertility is not an exogenous variable in the divorce equation, all these studies provide biased estimates of the effect of fertility on marital dissolution. There are two potential sources of bias. First, the couples might differ systematically in their observable characteristics by fertility, i.e. if characteristics like age at marriage, age, race, level of education and earnings differ between couples with children and childless couples, this might explain the observed marital dissolution differential by fertility. The second source of bias might be due to unobservable factors that affect both fertility and marital instability, in which case at least part of the observed relationship between them is spurious. The existence of any of the two biases would imply that households with children would behave differently from households with no children, independently of any true causal effect of fertility on divorce (selection bias problem).

Studies in the second category acknowledge the problem of endogeneity of fertility but their analysis is not always convincing. For instance, Becker et al. (1977) initially suggest the use of a simultaneous equations model to identify the causation between children and dissolution, but then they decide against this strategy by constructing a situation (they select women aged 40-55 whose fertility is already completed) that excludes causation running from marital (in)stability to fertility. Koo and Janowitz (1983) formulate, for married couples, a simultaneous model of fertility and marital dissolution, but then they estimate the two equations individually by single equation logit method, ignoring the issue of simultaneity previously addressed. Lillard and White (1993) use instrumental variables techniques to identify the simultaneous model of marital separation and fertility. However, the instrument chosen to identify the separation equation - the legal environment for divorce in the state of current residence - is weak because it need not affect separation, which is the outcome of interest- but only legal separation and divorce. Finally, Brien, Lillard and Stern (1999) propose a way to model endogenous investment in children in a model of cohabitation, marriage and divorce; unfortunately, they revert to exogenous investment in children as an element of the cost of divorce in their estimation because of computational costs.

This paper takes into account the problem of endogeneity of fertility but it uses the framework known as the *potential outcome approach* to identify and estimate the effect of interest. In

particular, the relationship between fertility and marital dissolution is formulated in a treatment-outcome framework similar to an experiment where the treatment is randomly assigned. The treatment of interest (fertility in this case) is defined in terms of potential marital outcomes for couples with children (treated). In particular the following question is explored: what would have been the marital outcome of a couple with children had they not had children? In order to answer to this question, this paper draws on matching methods developed in the statistics literature (Rubin 1977, 1979, Rosenbaum and Rubin 1983, Heckman et al. 1989, 1997, 1998) that exploit full information contained in observable covariates. The matching method provides a way to estimate treatment effects when controlled randomization is not possible and there are no convincing natural experiments which could substitute randomization. The main purpose of the method is to identify a systematic way to construct a correct sample counterpart for the missing information on the treated outcomes had they not been treated and to pair units in the groups of controls and treated.

The main findings indicate that parents do not divorce less in presence of children but they only postpone the decision to divorce until children get older. Furthermore, it is mainly the presence of children which discourages marital dissolution, while the number of children has only a small effect and in the direction of increasing the probability of divorce.

The remainder of the paper is organized as follows: Section 2 briefly summarizes the decision model of fertility and divorce presented in Vuri (2001). Section 3 introduces the potential outcome approach and it identifies the treatment effect under the causal effect model. Section 4 describes the matching approach for the binary case (*having children or not*). Section 5 extends the methodology to the multi-valued treatment case (*number of children*). Section 6 describes the data sets used (the Fertility and Marital History Supplement of Current Population Survey, June 1995), and the process of sample selection. In section 7, the results are presented. Finally, section 8 presents some concluding remarks and direction for further research.

2 The 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 cohabitations are seen as voluntary arrangements between two adults, formed to coordinate consumption and production activities, including the conception of children. When couples marry, they begin to acquire various "things" together, including a dwelling and its furnishings, shared interest to friends and so on. These are defined as "general" investments because they retain their value regardless of the couple's marital status. However, there are also other types of investments made by the couples called "marital-specific" because they belong to the couple rather than to either one of the partners separately (e.g. information on the partner's preferences, a division of labor inside and outside the household, sexual affinity and children). One immediate implication of this distinction between marital investments is the way they affect a couple's divorce probability because the accumulation of marital-specific capital raises the expected gain from remaining married and consequently discourages dissolution (Becker et al. 1977).

This is in particular the case for children because they represent the most important marital-specific investment of a couple during their marriage. Therefore, parenthood provides an important basis for marital stability and children greatly lower the risk of marital disruption (see Becker et al. 1977, Cherlin 1977, Becker 1991, Morgan and Rindfuss 1985). The presence of children may not only make the marriage more stable but it may also delay divorce[1] by increasing the gains from marriage and making it more costly than continuation in the marriage for two reasons: i) because of the anticipated complications attending a divorce action, such as problems with child custody, visitation plans, coparenting and single-parent problems; ii) because of the increasing awareness of the financial and psychological costs of divorce for children. Consequently, children appear to constitute financial, legal, and emotional[2] barriers to divorce.

However, causation also runs in the other direction, i.e. the arrival of children may be affected by the potential stability of the parents' marriage. In fact, a couple's "divorce inclination" may influence their decision to begin a family and their willingness to add children to an existing family. Therefore, 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).

In Vuri (2001) a simultaneous equations model of marriage status and fertility decisions which considers both directions of causality is presented. In particular, in the context of a model of marital-specific investment, it is shown that under some conditions, the 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. This framework leads (with some simplifications) to a simple estimable model described by the following two equations (for the complete derivation of the model see Vuri (2001, sections 4 and 5):

$$D_i = \beta C_i + \gamma \mathbf{X}_i + \epsilon_i \tag{1}$$

$$C_i = \delta \mathbf{X}_i + \nu_i \tag{2}$$

where D_i is the binary variable identifying whether the couple is observed to divorce (D_i =1) or to stay married (D_i =0); C_i represents an indicator of fertility;[3] $\mathbf{X}_i$ represents couple's demographic and social characteristics.

Equation 1 says that the decision to divorce is influenced by children in the household, by some observable characteristics $\mathbf{X}_i$ and by some unobservable factors ϵ_i. Equation 2 models the decision of a couple to have children, which depends on some observed characteristics $\mathbf{X}_i$ and some unobserved factors ν_i.

[1] Dissolution, if any, is measured as of the date husband and wife started to live separately, regardless of whether the legal formality of a divorce decree took place subsequently.

[2] I am referring in this case to a sort of "stigma" which is sometimes attached to persons who divorce when they have children, especially very young, which might discourage couples from divorcing.

[3] In sections 3 and 4, the fertility indicator is assumed to be binary, i.e. having children in different age groups; in section 5, the analysis is extended to the multivalued treatment "number of children in the household".

If fertility is exogenous to the divorce decision, then ordinary least square regression of the effect of fertility on marital dissolution yields an unbiased estimate of the treatment effect β in equation 1. However, fertility might be endogenous to the divorce decision if there is dependence between fertility C_i and the error term of the "divorce" relationship ϵ_i. The correlation between C_i and ϵ_i can arise for one of two not necessarily mutually exclusive reasons: (a) dependence between ϵ_i and ν_i,[4] or (b) dependence between $\mathbf{X}_i$ and ϵ_i.[5] The first case is referred as *selection on unobservables* (Heckman and Robb, 1985) and the second case as *selection on observables* (Rosenbaum and Rubin, 1983).

The methodology followed in this paper pursues the selection on observables approach and does not extend to selection on unobservables.[6] In what follows, the framework of the potential outcome approach to causality is introduced.

3 The potential-outcome approach

Using the terminology of the evaluation literature, let C_i denote a binary variable indicating treatment status "having children or not" ($C_i \in \{0,1\}$), $D_i(1)$ the potential marital outcome of a couple i under the treatment state "having children"($C_i = 1$), and $D_i(0)$ the potential marital outcome if the same couple i receives the alternative treatment (or no treatment) "having no children"($C_i = 0$). Thus, $D_i = C_i D_i(1) + (1 - C_i) D_i(0)$ is the observed marital outcome for the couple i. The individual treatment effect is $\beta_i = D_i(1) - D_i(0)$, which, however, is not observable since either $D_i(1)$ or $D_i(0)$ is missing. Alternatively, one might focus on the average effect of treatment on the treated couples (ATT henceforth):

$$\widehat{\beta}_{|C_i=1} = E(\beta_i \mid C_i = 1) = E[D_i(1) | C_i = 1] - E[D_i(0) | C_i = 1] \tag{3}$$

which implies comparing the marital outcome of a couple with children to the same couple in the counterfactual case of not having children, i.e. what would have been the marital outcome of a couple with children had not they had children. It is thus necessary that each couple is potentially exposable to any of the two treatments.[7]

While the first expectation $E[D_i(1) | C_i = 1]$ can be identified in the subsample of the treatment group, the counterfactual expectation $E[D_i(0) | C_i = 1]$ is not identifiable without invoking further assumptions. To overcome this problem, one has to rely on the untreated couples ($D_i(0)$) of the comparison group to obtain information on the counterfactual outcome of the treated in the no-treatment status. The replacement of $E[D_i(0) | C_i = 1]$ with $E[D_i(0) | C_i = 0]$ does not

[4] Because of the existence, for example, of an omitted variables like commitment to family.

[5] For example, because of the duration of marriage.

[6] The selection on unobservables hypothesis has been investigated in Vuri (2001), where IV techniques have been used to eliminate the bias due to omitted variables.

[7] Note that already at this stage the stable unit-treatment value assumption (SUTVA) has to be made. In our case, it requires that the marital outcome of a couple depends only on its fertility decision, not on the fertility decision, of other couples in the population, and that whether couples have children or not does not depend on the fertility decisions of others (no peers effect).

seem the right strategy since treated and untreated couples might differ in their characteristics determining the outcome if they themselves select into treatment.

An ideal randomized experiment would solve this problem because random assignment of couple to the treatment ensures that potential outcomes are independent of treatment status.[8] Hence, the treatment effect could consistently be estimated by the difference between the observed means of the outcome variable in the treatment group and in the no-treatment group. However, in this non-experimental setting, the choice of fertility is not likely to be random: fertility decision of a couple may depend on some observed characteristics which could also influence its marital outcome. For example, having a high level of education could affect both the decision to have children and at the same time the marital separation decision.

In this case, when randomized experiments are not available, other estimators have to be devised, relying on appropriate identifying assumptions.

In what follows, the approach used to construct a suitable comparison group, namely the matching method, and the identifying assumptions on which it is based, namely CIA, are described for the binary treatment case. In section 5 the approach is extended to the multivalued treatment case.

4 The matching approach

4.1 The Conditional Independence Assumption (CIA)

One approach to construct a correct sample counterpart for the missing information on the treated couples had they not been treated is based on statistical matching. Matching estimators try to re-establish the condition of an experiment when no such data is available by stratifying the sample of treated and untreated couples with respect to covariates X_i that rule both the selection into treatment and the outcome under study. Such a stratification eliminates selection bias provided all variables X_i are observed and balanced between treated and control group. In this case, each stratum (or cell) would represent a separate small randomized experiment and simple differences between treated and controls outcomes would provide an unbiased estimates of the treatment effect.

The matching method relies on the assumption that the relevant differences between any two couples, in terms of potential outcomes, are captured in their observable attributes. This underlying identifying assumption, called "conditional independence assumption"[9] (CIA henceforth) requires that, "conditional on observed attributes X_i, the distribution of the counterfactual outcome $D_i(0)$ in the treated group is the same as the (observed) distribution of $D_i(0)$ in the non-treated group" (Sianesi, 2001).[10] In other words, the outcomes of the non-treated are inde-

[8] Randomization implies that: $C_i \perp (D_i(0), D_i(1))$ and therefore: $E[D_i(0)|C_i = 1] = E[D_i(0)|C_i = 0] = E[D_i|C_i = 0]$

[9] Also "unconfoundedness", or "ignorable treatment assignment".

[10] This is actually the weaker version of CIA. The strong version (Rosenbaum and Rubin, 1983) asserts that the assignment to treatment C_i is unrelated to the pair of potential outcomes $(D_i(1), D_i(0))$, within subpopulations homogeneous in X_i. Formally: $C_i \perp (D_i(0), D_i(1)) \mid X_i$

pendent on the participation into treatment C_i, once one controls for the observable variables X_i. In symbols:

$$D_i(0) \perp C_i | X_i \tag{4}$$

It implies that, given X_i, the non-treated outcomes are what the treated outcomes would have been had they not been treated. This rules out the possibility that variables other than X_i, on which the analyst cannot condition, affect both $D_i(0)$ and C_i, i.e. there is no *selection on unobservables*.[11] Moreover, it assumes that $Pr(C_i = 0 \mid X_i = x) > 0$ for all x which guarantees that, with positive probability, there are untreated couples for each x.[12] From the previous two assumptions, it follows that $E(D_i(0) \mid X_i, C_i = 1) = E(D_i(0) \mid X_i, C_i = 0)$ (see Rosenbaum and Rubin 1983 for more details). The conditional mean response of the *treated* under no treatment for a given X can thus be estimated by the conditional mean response of the untreated under no treatment (the technique is simply to replace the unobserved outcomes of the treated had they not been treated with the outcome of non-participants with the same X_i characteristics, since they are statistically equivalent). In other words, the matched non-treated couples are used to measure how treated would have behaved, on average, had they not been treated.[13] [14]

4.2 The average treatment effect for the treated

Under the CIA, the average effect of treatment on the treated can be computed as follows:

However, since my objective is only the construction of the counterfactual $E(D_i(0) \mid X_i, C_i = 1)$ in equation 3, the weaker version of the CIA sufficies to identify the ATT.

[11] In other words, the CIA assumes that once two couples with the same history in terms of their observable characteristics have been found, it is likely that they have the same history also in terms of their unobservables. Consequently, being correlated, controlling for the observables it is possible to control for the unobservables at the same time.

[12] This implies to match couples only over the common support region of X_i where the treated and non-treated group overlap. Consequently, the ATT is computed only for those treated couples falling within the common support. The drawback of this selection is that if the treatment effect is heterogeneous across couples, restricting the sample of treated to the common support can change the parameter estimated.

[13] Note that under the "conditional assumption", it is not necessary to make assumptions about specific functional forms of outcome equations, decision process or distribution of unobservables.

[14] Note that the CIA is controversial because it is based on the assumption that the conditioning variables available to the econometricians are sufficiently rich to justify application of matching. In particular, the CIA requires that the set of the X_i's should contain all the variables that jointly influence the outcome with no-treatment $D_i(0)$ as well as the selection into the treatment (it is called a "data hungry" identification strategy by Heckman *et al.*, 1998). To justify the assumption, econometricians implicitly make conjectures about what variables enter in the decision set of couples, and how unobserved (by the analysts) relevant variables are related to observables.

$$\begin{aligned}\widehat{\beta}_{|C_i=1} &\equiv E\left[D_i\left(1\right)|C_i=1\right]-E\left[D_i\left(0\right)|C_i=1\right]= \\ &E_X\left\{E\left(D_i\left(1\right)\mid X_i,C_i=1\right)-E\left(D_i\left(0\right)\mid X_i,C_i=1\right)\mid C_i=1\right\}= \\ &\overset{CIA}{=} E_X\left\{E\left(D_i\left(1\right)\mid X_i,C_i=1\right)-E\left(D_i\left(0\right)\mid X_i,C_i=0\right)\mid C_i=1\right\}= \\ &E_X\left\{E\left(D_i\mid X_i,C_i=1\right)-E\left(D_i\mid X_i,C_i=0\right)\mid C_i=1\right\}\end{aligned} \tag{5}$$

The ATT is estimated by taking the difference of the outcomes in the two groups conditional on covariates and then averaging over the distribution of observable variables in the treated population $X_i \mid C_i = 1$.[15]

However, in a finite sample conditioning on X is problematic if the vector of observables is of high dimension. As the number of variables increases, the number of matching cells increases exponentially, and very often there will be cells containing either treated couples or control couples but not both, making the comparisons impossible.

Rubin (1977) and Rosenbaum and Rubin (1983) suggest to alternatively use the conditional probability to participate into the treatment $p(X_i) \equiv \Pr\left(C_i = 1 \mid X_i = x\right) = E\left(C_i \mid X_i\right)$, the *propensity score*, for purposes of stratifying the sample. They show that by definition treated and non-treated couples with the same value of the propensity score have the same distribution of the full vector of observables X_i. This is the so-called *balancing property* of the propensity score: $X_i \perp C_i \mid p\left(X_i\right)$. Furthermore, they demonstrate that if $D_i\left(0\right)$ is independent of C_i given X_i, $D_i\left(0\right)$ and C_i are also independent given $p(X_i)$. This implies that matching can be performed on $p\left(X_i\right)$ alone, thus reducing a potentially high dimensional matching problem to a one dimensional problem.

Matching treated and untreated couples with the same propensity scores and placing them into one cell means that the decision whether to participate or not is random in such a cell and the probability of participation in this cell equals the propensity score. Consequently the difference between the treatment and the non treatment average outcomes at any value of $p\left(X_i\right)$ is an unbiased estimate of the average treatment effect for the treated at that value of $p\left(X_i\right)$. Therefore, an unbiased estimate of the ATT can be obtained conditioning on $p\left(X_i\right)$, which is equal to exact matching on the $p\left(X_i\right)$. Formally:

$$\widehat{\beta}_{|C_i=1} = E_{p(X)}\left\{\left[E\left(D_i \mid C_i=1, p\left(X_i\right)\right)-E\left(D_i \mid C_i=0, p\left(X_i\right)\right)\right] \mid C_i=1\right\} \tag{6}$$

However, some drawbacks accompany this strategy. First, the propensity score itself has to be estimated (see Dehejia and Wahba (1998) for a description of the algorithm used to estimate the propensity score). Second, since it is a continuous variable, exact matches will rarely be achieved and a certain distance between treated and untreated couples has to be accepted. Several alternative and feasible procedures based on stratifying and matching (*nearest and radius*

[15] The regression equivalent of this procedure requires the inclusion of all the possible interactions between the observables X_i. The difference between regression and matching is the weighting scheme used to average estimates at different values of X.

methods of matching) on the basis of the estimated propensity score have been proposed in the literature to solve this problem (I refer the reader to the Appendix for a detailed description of the methods). In the next section, the extension of the matching methodology to the multivalued treatment case is described.

5 Estimation of the average causal effect with multi-valued treatment

Imbens (2000) and Lechner (1999) have proposed an extension of the propensity score methodology that allows for estimation of average causal effects with multi-valued treatments. The key insight of this method is that "for estimation of average causal effects it is not necessary to divide the population into subpopulations where causal comparisons are valid, as the propensity score does; it is sufficient to divide the population into subpopulations where average potential outcomes can be estimated" (Imbens, 2000: 706).

Here it is assumed that the treatment C_i can take values between 0 and K, i.e. $k = 0, 1, ..., K$. I am interested in average outcomes, $E\{D_i(k)\}$, for all values of k, and in particular in differences of the form $E\{D_i(k) - D_i(s)\}$, i.e. the average causal effect of exposing all units to treatment k rather than to treatment s. The key assumption, like in the binary case, is that adjusting for covariates solves the problem of drawing causal inferences. This is formalized by using a weak version of CIA in the multivalued treatment case. Let $T_i(k)$ be the indicator of receiving the treatment k:

$$\begin{aligned} \mathrm{T}_i(k) &= 1 \quad \text{if } \mathrm{C}_i = k \\ &= 0 \quad \text{otherwise} \end{aligned}$$

The weak version of CIA (also *weak unconfoundedness*) states that the assignment to treatment C_i is weakly unconfounded, given the covariates X_i if:

$$T_i(k) \perp D_i(k) \mid X_i \text{ for all } k = 0, ...K$$

Weak unconfoudedness requires only pairwise independence of the treatment with each of the potential outcomes. Furthermore, the independence of the potential outcome $D_i(k)$ with the treatment has to be only "local" at the treatment level of interest, i.e. with the indicator $T_i(k)$ rather than with the treatment level C_i. The important result is that under weak unconfoundedness the expected value of $D_i(k)$ can be estimated, by adjusting for X_i :

$$E\{D_i(k) \mid X_i\} = E\{D_i(k) \mid T_i(k) = 1, X_i\} = E\{D_i \mid C_i = k, X_i\}$$

Then, the average outcomes can be estimated by averaging the conditional means: $E\{D_i(k)\} = E[E\{D_i(k) \mid X_i\}]$.

Similarly to the binary case, it can be difficult to estimate $E\{D_i(k)\}$ when the dimension of X_i is large. To solve the problem, Imbens proposes the multivalued version of the propensity

score methodology. Firstly, Imbens defines the "generalized" propensity score (GPS), which is the conditional probability of receiving a particular level of the treatment given the covariates. In symbols:

$$r(k, x_i) \equiv \Pr(C_i = k \mid X_i = x_i) = E[T_i(k) \mid X_i = x_i]$$

Note that as in the binary case, GPS satisfies by definition the balancing property, i.e. $T_i(k) \perp X_i \mid r(k, X_i)$ for all $k = 0, ...K$.

Then, by using the same argument as in the binary treatment case, he proves that the CIA given the generalized propensity score $r(k, X_i)$ holds:

$$T_i(k) \perp D_i(k) \mid r(k, X_i) \quad \text{for all } k = 0, ...K$$

Consequently, the average outcomes is estimated by conditioning only on the generalized propensity score and the difference $E\{D_i(k) - D_i(s)\}$ can be easily computed for any k and s.

6 Data and measures

The empirical analysis of this paper is based on the *Current Population Survey* (CPS). The CPS is a monthly survey of about 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics.[16] In particular, I make use of the Fertility and Marital History Supplement File (June 1995), which provides retrospective information on fertility and marital history for women making possible it to identify some of the main characteristics of the women during their marriage (age at first marriage, duration of marriage, out of wedlock birth, number of children, age of each child etc.)

In order to build the final sample of analysis, I select women married only once.[17] The dependent variable is the indicator for marital status (equal to one if the woman divorces or separates in 1995 or before, 0 if the woman is still married in 1995). The covariates of interest are the indicator of the *Number of children between 0 and 18 years old (*i.e. children potentially still at home)*, Number of children between 0 and 6 years old, Number of children between 6 and 18 years old,* and dummy variables for *Having children between 0 and 18* , *Having the youngest child between 0 and 6, Having the youngest child between 6 and 18.*[18] Each record describes personal characteristics, like duration of marriage, age at marriage, out of wedlock birth, highest level of education, race, and Hispanic origin.

After restricting the sample to women with complete records in the critical variables, 25914 records remain. Tables 1 shows descriptive statistics (mean and standard deviation) of the covariates of interest included in the regressions for the selected sample. The table shows that

[16] For more information, see the web page http://www.bls.census.gov from which the data were downloaded.

[17] I do not consider higher order marriages because the factors affecting divorce in this cases can be different from the factors influencing the divorce of people married for the first time. Furthermore, only women are considered as representative of the couple because for the divorced husbands fertility history is not recorded.

[18] These variables are collected in 1995 for women still married in 1995, and at the date of divorce for women who divorced.

around 20% of the marriages end in divorce; however, the average duration of marriage is still high (17.5 years) but it is due to the fact that both divorced women and women whose marriage lasted long are considered. Most of the women have a medium level of education (63%), they marry quite young (22.1 years), around 22% have a child before marriage. Finally, most of the sample is constituted by white women. Table 2 shows summary statistics of the covariates for the three groups of treated (couples with children aged 0-18, aged 0-6 and aged 6-18 respectively) and controls (couples with no children in the three age groups), and t-statistics from the test of equality of means between them are reported. The tests show that there are relevant differences between the two groups in the three cases under study in terms of most of the explanatory variables (7 variables out of 11 for the treatment *Having children between 0 and 18* , 9 out of 11 for the treatment *Having the youngest child between 0 and 6,* ad 8 out of 11 for the treatment *Having the youngest child between 6 and 18).* The existence of such differences highlights the need for the careful statistical adjustment procedures described in the previous sections.

The variables selected in this paper are (as far as possible) the same considered in the literature on this topic. In particular, previous studies have usually included husband-wife characteristics at the time of marriage (or at the time of interview) like education, age, marriage duration, earnings, previous cohabitation, pre-marital births, and traits that for most individuals do not vary over time like religion, race etc. (see Lecher 2988, Becker et al. 1977, Brien, Lillard and Stern 1999, Lillard and Waite 1993, Ermish and Francesconi 1996).

For this reason, it can be reasonably assumed that the attributes considered in this analysis contain relevant observable information influencing both the marital outcome and fertility decision (of course, it would have been preferable to have also information on earnings, but education can be used as a reasonable proxy for future earnings). Therefore, the conditional independence assumption is assumed to be valid for the remainder of this paper.[19]

7 Estimation results

In section 7.1, estimates of the effect of several measures of fertility on marital dissolution outcome obtained through parametric methods, namely OLS and probit, are presented.[20] I present these overall estimates to facilitate comparison of my results with those of previous authors and to provide a benchmark for my later results. The analysis is performed both for the binary treatment and for the multivalued treatment.

[19] Of course, there may be substantial arguments claiming that this is not true. For example, if one believes that there are additional unobserved factors correlated with outcomes and selection into treatment, not captured when we condition on the observables, then, of course, this invalidates the CIA and the following analysis. Moreover, the presence of infertile women could also invalidate the CIA, because any woman has to be potentially exposable to both the treatments. However, it can be reasonably assumed that the proportion of infertile women in the sample is negligible.

[20] Probit estimation takes into account the binary nature of the outcome variable.

7.1 OLS and probit estimates

Firstly, the binary measures of fertility "*having children aged 0-18*"is considered. The OLS-estimates of the effects of having at least one child still at home (aged 0-18) on the probability of marital dissolution is reported in table 3 (row1), controlling for the vector of observed variables, indicated by X_i in equation (7) and listed in section 6. The estimated coefficient on the fertility binary variable "*having children aged 0-18*" is negative and equal to -0.124, which implies that having at least one child still at home has a negative effect on the dissolution rate (it decreases the probability of divorce). However the data can provide further information on the effect of having children in pre-school age or older children by disentangling the treatment "*having children aged 0-18*" into two measures of fertility, i.e."*having the youngest child aged 0-6*" and "*having the youngest child aged 6-18*". The results in table 3 (rows 2 and 3) show that the negative effect previously estimated is almost entirely due to the effect of having young children (aged 0-6) on dissolution (-0.129), while having older children (aged 6-18) increases significantly the dissolution probability but the effect is much smaller in size (0.011). The same result is obtained from probit estimation (columns 2).[21]

In rows 4-6 of table 3, the OLS and probit estimates for the multivalued case (when the treatments are "*number of children aged 0-18*", "*number of children aged 0-6*","*number of children aged 6-18*") are reported. These results are similar to the ones for the binary case, but smaller in size. In particular, an additional child in each of the three groups of interest has an effect respectively of -1.5%, -3.8% and 2.3% on marital dissolution. From these results it could be argued that mainly the presence of children in the three groups matters, while their numerosity has a much smaller effect on marital dissolution.

However, as already pointed out, OLS (or probit) estimates can be biased because of the self-selection problem and the potential correlation of fertility with some observable characteristics that make these estimates biased. Therefore, I turn to the propensity score matching estimators which provide unbiased estimates of the causal effect of fertility on divorce.

7.2 Results using the propensity scores

This section is organized in the following way: the first part focuses on the analysis of the binary treatment case, in which the estimation of the propensity score, and the results from the stratification and matching procedures are presented; the second part is devoted to the multi-valued treatment results.

7.2.1 The dichotomous treatment case

Estimating the propensity score The first step in the implementation of this methodology is to estimate the propensity score for the three treatments under study. In general, any standard probability model can be used to estimate the propensity score. For example, $\Pr\{C_i = 1 \mid X_i\} = F(h(X_i))$, where F(.) is the normal or the logistic cumulative distribution and $h(X_i)$ is a

[21] Marginal effects of probit estimation are reported in all the tables in order to be comparable with the OLS estimates.

function of covariates with linear and higher order terms. In this paper, the propensity score for the three treatments of interest is estimated using a probit model and following the algorithm proposed by Dehejia and Wahba (1998), which suggest to group the observations into blocks defined on the estimated propensity score and to test the balancing property for the score and the covariates. In my case, I start with five blocks based on the quintiles of the estimated propensity score for the treated, and then I test whether the means of the score for the two groups are statistically different. Finer blocks are built until the test is satisfied for all of them and 12 blocks are identified for the treatment "*having children aged 0-18*", 13 blocks for the treatment "*having the youngest child between 6-18 years old*", and 11 blocks for the treatment "*having the youngest child between 0-6 years old*". Once the balance is achieved for the score, also the distributions of covariates X_i between the two groups should be identical for the balancing property. I provide an example of it by testing for equality of means between the treated and the control groups for each of the eleven variables in X_i, within each block and for each treatment. In almost all cases I find equality of means of the X_i at the 5% confidence level, and none of the covariates does systematically fail the test in all the blocks. Remember that when the same test was performed on the whole sets of control and treated units, rather then within each stratum, I rejected equal means for seven out of eleven variables used in the regression for the treatment "*having children 0-18",* nine out of eleven for the treatment *"having the youngest child 0-6"* , and eight out of eleven variables for the treatment "*having the youngest child 6-18*" (see Table 2). Figure 1 plots the histograms of the estimated propensity scores for the three treatments. Note that they do not include the controls whose estimated propensity score is less than the minimum or more than the maximum estimated propensity score for the treated units. This selection is necessary in order to guarantee that the treated and the control units lie on a common support. The figure reveals that there is a good overlap in terms of the propensity score in each block (especially for the two treatments "*having the youngest child 0-6"* and *"having the youngest child 6-18*"), while in the extreme bins there is only a limited overlap, as expected, because the number of treated units increases and the number of control units decreases at high values of the propensity score. However, this does not generate bias in the estimates as long as the balancing property is satisfied; this ensures that the treated couples in each block are observationally identical to the controls in the same block and only by chance does the treatment status differ in the two groups.

The next step consists in estimating the ATT using equation 6, where the ATT is computed as the difference between the treatment and the control average outcomes at any value of $p(X_i)$. However, as mentioned in section 4, the exact matching on $p(X_i)$, implicit in this strategy, is unfeasible in practice and thus stratifying and matching (*nearest and radius methods of matching*) on the basis of the estimated propensity score are used.

Estimating the treatment effect

Blocking (stratification) estimator The stratification estimator relies on the same division into strata defined in section 7.2.1, where the covariates are balanced across treated and control

couples by construction. Then, within each block, the difference between the average outcomes of the treated and the controls is computed. The ATT of interest is finally obtained as an average of the ATT of each block with weights given by the distribution of treated units across blocks (see the Statistical Appendix for the formula of the stratification estimator).

The stratification estimates show a negative effect of fertility on divorce for the cases "*having children aged 0-18*" and "*having the youngest child aged 0-6*" , the second effect being larger than the first one (see table 4). In fact, having children at home decreases the probability of divorce by 8.8 percentage points, while having the youngest child between 0 and 6 years old decreases this probability by 10.2 percentage points. Instead, having the youngest child aged 6-18 increases marital dissolution probability by 0.6 percentage points**.**

An alternative is the linear and probit regression adjustments, which consist in regressing the marital outcome indicator on the fertility indicator and all the covariates used to estimate the propensity score; this should help eliminating the remaining within-block differences in the covariates. The results are similar to the estimates reported in table 4 (and therefore they are not reported) providing further evidence that the covariates are well balanced.[22]

Matching estimator This method consists of taking each treated unit and searching for the control unit with the closest propensity score, i.e. the *best match*. The method is usually applied with replacement implying that a control unit can be a best match for more than one treated unit. Then the ATT of interest is obtained by averaging the differences between the outcome of the treated unit and the outcome(s) of the matched control unit(s). However, it is obvious that some of these matches are fairly poor because the nearest matched unit is too far to be considered as a valid match. The *radius method* of matching offers a solution to this problem because it consists in matching each treated to the control couple(s) whose propensity score is within a $\delta-$radius chosen by the researcher.[23] In this way only higher quality matches are selected, even if it has the disadvantage of reducing the sample size (see the Statistical Appendix for the formula of the matching estimators).

In table 4, the results of the matching estimators computed using both the *nearest matching method* and the *radius method* are reported for the three treatments of interest. For the last method, three different measures of the δ-radius are chosen in order to check the robustness of the estimates to this choice. Note that, in the δ-radius method, not only the more distant controls are discarded, like in the nearest matching method, but also the treated units for which a match within the δ chosen could not be found. Consequently the lower is the radius chosen, the smaller is the number of remaining units.

Like for the stratification estimator, the treatment effect can be either estimated as a difference in means in marital outcomes across these pairs of treated and matched control units, or as a linear regression of marital outcomes on treatment and covariates on the balanced sample. Since the results are similar across the two methods, providing further evidence that controlling for covariates does not alter the estimates significantly, only the first ones are reported.

[22] The results of linear regression in matching are available on request from the author.

[23] The choice of δ depends on the willingness of the researcher to select more accurate matches because lower δ implies selecting higher quality matches.

In particular, the matching estimates provide evidence of a negative effect of *having children 0-18* and of *having the youngest child aged 0-6* on marital dissolution (and significant in all the cases). The first effect ranges from -0.133 for the nearest method to -0.039 for radius method with $\delta = 0.0001$, the second effect ranges from -0.155 for the nearest method to -0.082 for radius method with $\delta = 0.001$. The estimate of the effect of *having the youngest child aged 6-18* on the contrary gives mixed results, because it is positive for the nearest matching method (0.017) and negative for the radius method (from -0.0011 for $\delta = 0.001$ to -0.002 for $\delta = 0.0005$). However, the results from the radius method are never significant and therefore only the first estimate should be considered reliable.

In conclusion, matching methods and stratification method yield similar results to the parametric methods (OLS and Probit) reported in section 7.1 in terms of direction of the effect, but they are bigger in size; it provides evidence that OLS estimates tend to underestimate the true effect of fertility on divorce. The estimates obtained through matching methods support the idea that having young children only delay parental's decision of divorce until children get older or also that when children reach the school age, additional problems (financial, organizational, etc.) could arise and generate distress inside the couple.

7.2.2 The multi-valued treatment case

Table *5* shows the estimates in the multi-valued treatment cases, respectively "*number of children aged 0-18*", "*number of children aged 0-6*" and "*number of children aged 6-18*". It is important to note that in this case not only the average effect of having an additional child on the probability of divorce can be computed (as for the OLS and Probit estimates, see table 3), but the effect of each additional child on marital disruption can be analyzed, i.e. the effect of going from 0 to 1 child, from 1 to 2 children, from 2 to 3, from 3 to 4 children and so on.[24] In order to compute the generalized propensity score an ordered probit of each of the three treatments of interest on the covariates X_i is computed.

The estimates show that there is a positive effect of the treatment "*number of children aged 0-18*" on divorce for going from 0 to 1 child and from two to four (increases marital dissolution probability), and negative in the other cases (discourages marital dissolution). By computing a weighted average of these effects, the evidence shows that having an additional child in the range 0-18 years old increases on average the probability of marital dissolution by 1.2 percentage point. Like in the binary case, I first look at the results for the treatment "*number of children aged 0-6*", and it turns out that the effects on marital disruption is strongly positive for the first child and then it turns out to be negative for children of higher order. This surprisingly strong positive effect for the first child can be attributed to the fact that many couples have the first child out of wedlock and then they marry. This kind of marriages are at higher risk of disruption and the effect could be more visible for couples married for less years and with young children. The average effect in this case is 0.05 percentage points. Finally, the estimates of the treatment

[24] The women with more than 5 children were too rare to be considered separately; consequently, all the women with more than five children have been grouped. For the case *number of children 0-6* the last group is constituted by women having 4 or more than 4 children.

"*number of children aged 6-18"* on marital dissolution are positive but not significant in many cases. The average effect shows that having an additional young child increase significantly the probability of divorce by 0.7 percentage points.

Note that in two cases out of three ("*number of children aged 0-6*" and "*number of children aged 0-18")* these estimates are of different sign with respect to the OLS estimates (which were negative). This would imply that OLS method tends to underestimate the true effect. Furthermore, these estimates are different from what obtained in the binary case, i.e. they are positive and small in size, suggesting that a high number of children might generate distress in the family, consequently increasing the probability of marital dissolution. However, this effect is very small.

8 Conclusion

In this paper, I have shown how to estimate the treatment effect of fertility on marital dissolution in presence of non random assignment using propensity score methods. In particular, I have analyzed the effects of three binary treatments "*having children aged 0-18*", "*having the youngest child aged 0-6*", and "*having the youngest child aged 6-18*" on marital dissolution, by using stratification and matching techniques, and the effects of the multivalued treatments "*number of children aged 0-18*", "*number of children aged 0-6*", and "*number of children aged 6-18*" on marital dissolution. The empirical analysis strengthens the evidence that parents do not divorce less in presence of children but they only postpone the decision to divorce until children get older. Another possible explanation is that additional problems could arise inside the couple when children get older. In addition, the estimates for the multivalued case show that the number of children has a positive bt small effect on marital dissolution. Overall, the results suggest that the presence of young children seems to enforce the marriage while the number of children in each age groups does not seem to have a relevant impact on marital dissolution.

As mentioned in section 2, the propensity score method solves the problem of selection on observables but not the one of selection on unobservables. The most recent research is pointing to the direction of a mixture of the two strategies, i.e. the selection of comparable individuals based on matching method to control for selection on observables and then application of fixed effect procedure (or difference in difference method) to control for selection on unobservables. Investigation on this recent proposal represents work for future research.

References

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.

Angrist J.D. and Krueger A.B., (1998), "*Empirical Strategies in Labor Economics*", Handbook of Labor Economics (2000).

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.

Becker G.S., (1991); A Treatise on the Family; Harvard University Press.

Blundell R. and Costa Dias M. (2000), "*Evaluation Methods for Non-Experimental Data*", mimeo UCL

Brien M.J., Lillard L.A. and Stern S., (1999), "*Cohabitation, Marriage, and Divorce in a Model of Match Quality*", mimeo, University of Virginia.

Dehejia R and Wahba S. (1998), "*Propensity Score Matching Methods for Non-Experimental Causal Studies*", NBER Working Paper n° 6829

Dehejia R and Wahba S. (1999), "*Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs*", JASA: 1053-1062

Ermish J. and Francesconi M. (1996), "Pertnership Formation and Dissolution in Great Britain", WP 96-10 of the ESRC, Colchester, University of Essex.

Green W.H., (1987), Econometric Analysis, 3rd. edition, Prentice Hall.

Heckman J. and Hotz V.J. (1989), "*Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: the Case of Manpower Training*", JASA December: 862-880.

Heckman J., Ichimura H, and Todd P. (1998), "*Matching as an Econometric Evaluation Estimator*", Review of Economic Studies 65: 261-294

Heckman J., Ichimura H, and Todd P. (1997), "*Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme*", Review of Economic Studies 64: 605-654

Heckman J., Ichimura H, Smith J.A., and Todd P. (1998), "*Characterising Selection Bias Using Experimental Data*", Econometrica 66.

Koo H and Janowitz B., (1983), "*Interrelationship between fertility and marital dissolution: results of a simultaneous logit model*", Demography 20 (2): 129-145.

Imbens G., (2000), "*The Role of Propensity Score in Estimating Dose Response Functions*", Biometrika 87.

Imbens G., Rubin D.B., and Sacerdote B, (1999), "*Estimating the Effect of Unearned Income on Labor Supply, Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players*", NBER Technical Working Paper Series n°7001.

Lechner M. (2001), "*Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies*", University of St Gallen, mimeo.

Lillard L.A. and Waite L.J. (1993), "*A Joint Model of Marital Childbearing and Marital Disruption*", Demography 30: 653-681.

Manski C.F., Sandefur G.D., McLanahan S. and Powers D., (1992), "*Alternative Estimates of the Effect of Family Structure During Adolescence on High School Graduation*", JASA 87: 25-37.

Ono H., (1998), "*Husbands' and Wives' Resources and Marital Dissolution*", Journal of Marriage and the Family 60: 674-689.

Peters E., (1986), "*Marriage and Divorce: Informational Constraints and Private Contracting*", The American Economic Review 76: 437-454.

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.

———, (1985a), "*Reducing Bias in Observational Studies Using the Subclassification on the Propensity Scores*", Journal of the American Statistical Association, vol.79: 516-524.

———, (1985b), "*Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity*", American Statistician, vol.39(1): 33-38.

Rubin D.„ (1979), "*Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observation Studies*", Journal of the American Statistical Association, vol. 74: 318-328.

Sianesi B. (2001), "*An Evaluation of the Active Labor Market Programs in Sweden*", mimeo UCL

Thornton A, (1977), "*Children and marital stability*", Journal of Marriage and the Family 39: 531-540.

Vuri D. (2001), "*Fertility and Divorce*", European University Institute Working Paper n°5.

Waite L. and Lillard L., (1991), "*Children and marital disruption*", American Journal of Sociology 96: 930-53.

Waite L., Haggstrom G. and Kanouse D., (1985), "*Parenthood and marital stability*", American Sociological Review 50: 850-57.

Weiss, Y and Willis, J.R., (1985), "*Match Quality, New Information and Marital Dissolution*", Journal of Labor Economics 15: S293-S329.

9 Statistical Appendix

9.1 The blocking estimators

The strata are chosen so that the balancing property is satisfied. Then, within these blocks indexed by q, the average difference in marital status between the treatment and the control units is computed:

$$\widehat{\beta}_q = \frac{\sum_{i\in I(q),C_i=1} D_i}{\sum_{i\in I(q)} C_i} - \frac{\sum_{i\in I(q),C_i=0} D_i}{\sum_{i\in I(q)} (1-C_i)} = \frac{\sum_{i\in I(q),C_i=1} D_i}{N_q^T} - \frac{\sum_{i\in I(q),C_i=0} D_i}{N_q^C}$$

where N_q^T and N_q^C are the numbers of treated and controls in block q, and $I(q)$ is the indicator function for the woman i being in the block q; then, in order to extend this result to the entire population of treated, the weighted average of these differences is computed:

$$\widehat{\beta}_{|C_i=1} = \sum_{q=1}^{Q} \widehat{\beta}_q \frac{\sum_{i\in I(q)} C_i}{\sum_{\forall i} C_i}$$

where the weights in each stratum are the fraction of treated units in each block. The standard errors for this estimator reported in tables 6-8 have been computed by boot-strapping with 200 repetitions.

An alternative is a linear regression or covariance adjustment techniques within each block. In this case, the treatment effect β_q is obtained by regressing D_i on the treatment indicator of fertility and other covariates *within* each block, and then computing a weighted average exactly as outlined before (the weights are identical).The advantage of the "regression in matching" is that controlling again on the covariates X_i should help to eliminate the remaining within block-differences, although the results should no change when the covariates are well balanced.

9.2 The matching estimator

In the *nearest-match method,* each treated unit is matched to the control unit(s) with the closest propensity score. In symbols, the treated woman i is matched to that non-treated woman j such that:

$$p(X_i) - p(X_j) = \min_{k\in\{C_i=0\}} \{ \mid p(X_i) - p(X_k) \mid \}$$

None of the treated women is discarded in the nearest-match method because it is always possible to find a matched control even if it is far away from the treated woman.

The *radius method* of matching consists in matching each treated to the control women whose propensity score is within a $\delta-$radius chosen by the researcher. In symbols:

$$\delta > p(X_i) - p(X_j) = \min_{k\in\{C_i=0\}} \{ \mid p(X_i) - p(X_k) \mid \}$$

If a treated unit has no control couples within a $\delta-$radius, this unit is discarded. Hence, switching from the nearest-match to the radius match one improves the quality of the matches but ends up using less observations and thus generate less precise estimates.

The average difference in marital outcomes of the sub-group of treated and the sub-group of matched comparisons is used to calculate the effect of having children *versus* not having children

on marital dissolution. Formally, the matching estimator is:

$$\widehat{\beta}^{M}_{|C_i=1} = \frac{1}{N^T}\left[\sum_{i\in T} D_i - \sum_{i\in C}\omega_i^C D_i\right] \tag{7}$$

where T and C denote the sets of treated and matched control units respectively and ω_i^C is the number of times a particular control $i \in C$ is used in the matching with a treated unit. Therefore, the average treatment effect is simply given by the average of the outcome in case of treatment minus the weighted average of the outcome in case of no treatment, with appropriate weights for repeated observations.[25] The standard error for this estimator reported in tables 6-8 have been computed by boot-strapping with 200 repetitions.

As for the stratification estimator, an alternative is to regress, over the sample of pairs, the divorce outcome on the treatment indicator of fertility and the covariates, with appropriate weights for repeated units (the logic for computing the weights is the same as described above). Note that, while in the stratification method the regression was simply an OLS estimation within blocks, in the matching method a weighted least square regression (WLS) is performed, where the weights are one for the treated and the number of times each control units is used in the matching for the controls (see equation (7)).

[25] Note that N^T is equal to the number of all treated units in the nearest match methods, and to the number of the treated units for whom at least one matched control could be found in the radius method.

Table 1: Descriptive statistics of observable covariates for all the women(sample size 25914)

Variable	All women	
	mean	st.dev
divorce rate	0.18	0.38
# children 0-18	1.11	1.21
# children 0-6	0.42	0.76
# children 6-18	1.59	1.47
having children 0-18	0.55	0.49
having children 0-6	0.27	0.44
having children 6-18	0.28	0.45
Duration of marriage	17.5	12.3
No Diploma	0.12	0.33
Some Schooling	0.63	0.48
Degree or more	0.24	0.43
Out of Wedlock birth	0.22	0.41
Age at marriage	22.1	4.08
White	0.84	0.36
Black	0.09	0.28
Others	0.07	0.26
Hispanic origin	0.08	0.27

Data legend: The three variables "Number of children between 0-18, 0-6 and 6-18" range from 0 to 5 where 5 means having five or more than five children; Duration: duration of the marriage; Highest level of education ranging from 31 (less than 1st grade) to 46 (Doctorate degree) has been divided into three dummies: No Diploma (from 31 "Less than 1st Grade" to 38) "12th Grade No Diploma", Some Schooling (from 39 "GED" to 42 "Associate Deg") and Degree or more(from 43 "Bachelor" to 46 "PhD") ; White: 1 if woman of white race; Black: 1 if woman of black race; Others: 1 for all others race; Out of Wedlock birth=1 if first child was born before marriage; hispanic: 1 if Hispanic origin.

Table 2: Descriptive statistics of observable covariates by children aged 0-18, 0-6 and 6-18.

Variable	kids 0-18	no kids0-18	kids 0-6	no kids 0-6	kids 6-18	no kids 6-18
divorce rate	0.20 (0.4)	0.16(0.37)	0.20(0.4)	0.17(0.37)	0.19(0.39)	0.18(0.38)
		[7.38]		[6.8]		[1.3]
Duration	13.24(7.7)	22.96(14.7)	8.3(5.0)	21.8(12.4)	18.1(17.9)	17.2(17.1)
		[68.4]		[84.3]		[4.9]
Degree or more	0.24(0.43)	0.23(0.42)	0.27(0.44)	0.23(0.42)	0.23(0.42)	0.24(0.43)
		[1.48]		[4.4]		[2.78]
Some schooling	0.64(0.5)	0.63(0.5)	0.62(0.48)	0.64(0.48)	0.65(.47)	0.62(0.48)
		[1.6]		[3.07]		[4.8]
No Diploma	0.11(0.32)	0.13(0.34)	0.11(0.32)	0.12(0.33)	0.11(0.31)	0.13(0.33)
		[4.3]		[1.2]		[3.5]
Age at marriage	22.1(5.3)	22.2(5.3)	22.4(4.2)	22.0(5.0)	21.8(4.4)	22.8(4.9)
		[0.29]		[6.4]		[6.7]
Out of Wedlock	0.27(0.4)	0.14(0.3)	0.27(0.44)	0.19(0.39)	0.26(.44)	0.19(0.4)
		[24.4]		[15.1]		[11.6]
White	0.83(0.4)	0.85(0.3)	0.83(0.37)	0.86(0.36)	0.83(.37)	0.84(0.36)
		[4.8]		[2.7]		[2.6]
Black	0.09(0.3)	0.08(0.3)	0.08(0.27)	0.09(0.28)	0.09(0.29)	0.08(0.28)
		[1.1]		[1.6]		[2.9]
Other	0.08(0.26)	0.06(0.24)	0.08(0.27)	0.06(0.24)	0.07(0.25)	0.07(0.26)
		[5.6]		[5.7]		[0.5]
Hispanic	0.09(0.3)	0.06(0.2)	0.11(0.31)	0.06(0.25)	0.08(0.27)	0.08(0.27)
		[11.6]		[12.9]		[0.04]

Notes: Standard deviation in parentheses. In square brakets, the t-statistics of the difference in means between women with and without children in the three range groups: 0-18, 0-6 and 6-18

Table 3: Parametric Estimates of the effect of different measures of fertility on marital dissolution

		OLS	Probit
		(1)	(2)
(1)binary treatment[a]	having children 0-18	-0.124	-0.088
		(0.007)	(0.007)
	having children 0-6	-0.129	-0.123
		(0.006)	(0.006)
	having children 6-18	0.011	0.041
		(0.005)	(0.005)
(2) multi-valued treatment[b]	# children 0-18	-0.015	-0.005
		(0.002)	(0.002)
	# children 0-6	-0.038	-0.040
		(0.003)	(0.003)
	# children 6-18	0.025	0.033
		(0.002)	(0.002)

Notes: In column 2 (probit) marginal effects are reported .

[a]Least square regression: marital dissolution dummy on a costant, a fertility treatment indicator, duration of marriage, age at marriage, two dummies for education, black dummy, other race dummy, hispanic origin dummy, out of wedlock dummy.

[b]Least square regression: divorce dummy on a fertility multi-valued treatment indicator and on the same covariates as in [a].

Table 4: Propensity score estimates of the effect of the presence of children between 0 and 18 years old, between 0 and 6 years old, and between 6 and 18 years old on marital dissolution (ATT)

	Having children 0-18	Having children 0-6	Having children 6-18
Stratification:			
based on quintiles	-0.088	-0.102	0.006
	(0.033)	(0.04)	(0.004)
Matching:			
Nearest Match	-0.133	-0.155	0.017
	(0.008)	(0.007)	(0.006)
Radius: δ<0.0001	-0.039	-0.093	-0.0015
	(0.005)	(0.007)	(0.007)
Radius: δ<0.0005	-0.048	-0.085	-0.0020
	(0.004)	(0.007)	(0.006)
Radius: δ<0.001	-0.065	-0.083	-0.0011
	(0.005)	(0.008)	(0.007)

Notes: Coefficients on the binary variables "Having children between 0 and 18 years old", "Having children between 0 and 6 years old"and "Having children between 6 and 18 years old" are reported. Boot-strapped standard errors in parentheses.

Propensity scores are estimated using the probit model, with the following specification:

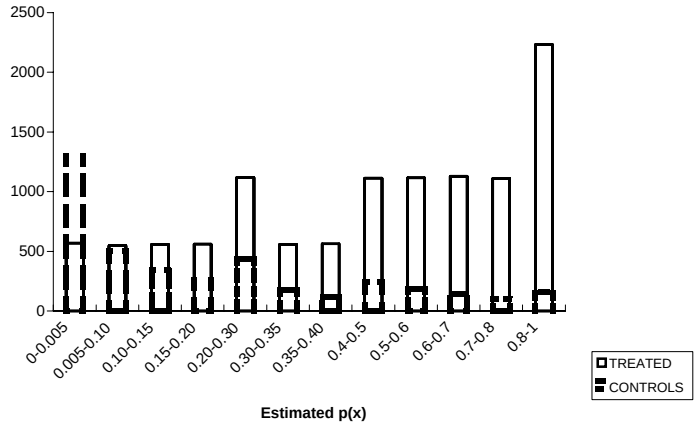
$\Pr(C_i=1)=F$(marital duration, age at marriage, 2 dummies for education, black race dummy ,other race dummy, hispanic origin dummy, out of wedlock birth

Table 5: Propensity score estimates of the effect of number of children aged 0-18, aged 0-6 and aged 6-18 on marital dissolution

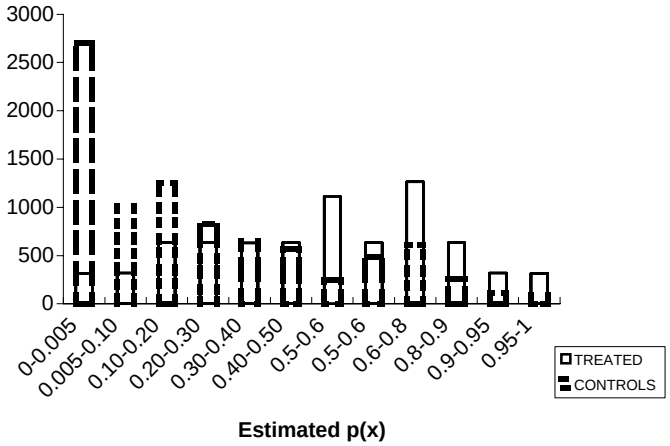
	# children 0-18	# children 0-6[a]	# children 6-18
From 0 to 1 child	0.03	0.124	0.002
	(0.005)	(0.006)	(0.006)
From 1 to 2 children	-0.002	-0.056	0.011
	(0.001)	(0.001)	(0.006)
From 2 to 3 children	0.013	-0.067	0.003
	(0.0002)	(0.003)	(0.005)
From 3 to 4 children	0.004	-0.088	0.006
	(0.001)	(0.003)	(0.002)
From 4 to 5(+) children	-0.001	-	0.009
	(0.001)	-	(0.003)
Average effect	0.012	0.05	0.007
	(0.002)	(0.003)	(0.002)

Note: Bootstrapped standard errors in parentheses. The covariates X included in the regression are the same listed in table 3. [a] For the treatment "number of children aged 0-6", the maximum number of children considered is 4 or more than 4. For the treatments "number of children aged 0-18" and "number of children aged 6-18", the maximum number of children considered is 5 or more than 5.

Histogram of the Estimated Propensity Score for the Treated and the Controls (having children aged 0-18)



Histogram of the Estimated Propensity Score for the Treated and the Controls (having the youngest children aged between 0-6 years)



Histogram of the Estimated Propensity Score for the Treated and the Controls (having the youngest child aged 6-18)