

The Role of Family Orientations in Shaping the Effect of Fertility on Subjective Well-being: A Propensity Score Matching Approach

Nicoletta Balbo¹ · Bruno Arpino²

Published online: 15 June 2016

© Population Association of America 2016

Abstract This article investigates whether and how having a child impacts an individual's subjective well-being, while taking into account heterogeneity in family attitudes. People with different family orientations have different values, gender attitudes, preferences toward career and family, and expectations about how childbearing can affect their subjective well-being. These differences impact fertility decisions and the effect of parenthood on an individual's life satisfaction. We define three groups of people based on their family orientations: Traditional, Mixed, and Modern. Applying propensity score matching on longitudinal data (British Household Panel Survey), we create groups of individuals with very similar socioeconomic characteristics and family orientations before childbearing. We then compare those who have one child with those who are childless, and those who have two children with those who have only one child. We show that parents are significantly more satisfied than nonparents, and this effect is stronger among men than among women. For men, we do not find significant differences across family orientations groups in the effect of the birth of the first child on life satisfaction. Among women, only Traditional mothers seem to be more satisfied than their childless counterparts. Women who have a second child are never more satisfied than those who have only one child, regardless of their family orientations. Traditional and Mixed men experience a gain in life satisfaction when they have a second child, but this effect is not found for Modern men.

Keywords Life satisfaction · Fertility · Family orientations · Propensity score matching

✉ Nicoletta Balbo
balbo.nicoletta@unibocconi.it

¹ Dondena Centre, Bocconi University, Via Guglielmo Rontgen 1, 20136 Milan, Italy

² Pompeu Fabra University, Barcelona, Spain

Introduction

Recent years have witnessed a growing recognition of the importance of the link between fertility and subjective well-being (e.g., Aassve et al. 2012; Billari 2009; Margolis and Myrskylä 2011; Myrskylä and Margolis 2014). This interest stems from the fact that in contemporary advanced societies, where contraceptive use is widespread, the choice of having children is largely determined by an individual's subjective well-being, both present and prospective. Although having at least one child is still sometimes considered normative or at least expected from couples, the link between fertility and subjective well-being has been identified as an important missing link in explaining fertility trends and differences across space in advanced societies (Billari 2009).

Cross-sectional research has so far shown no consensus about the effect of fertility on life satisfaction. Some studies seem to imply that in advanced societies, children are detrimental to an individual's subjective well-being (SWB) because they reduce marital well-being (McLanahan and Adams 1987) or increase the stress associated with financial responsibility (Zimmermann and Easterlin 2006); other studies emphasize a positive association between childbearing and SWB. The value of children theory (Friedman et al. 1994; Hoffman and Hoffman 1973; Hoffman and Manis 1979) envisions having children as instrumental for maximizing individual utility as expressed by the combination of physical well-being and social esteem. In this sense, having children when fertility control is available positively contributes to individual well-being. In line with this approach, needs theories also stress the advantages associated with having children, with parenthood bringing emotional rewards (Veenhoven 1996). Also emphasizing a positive association between fertility and SWB are studies claiming that children help parents increase their social capital and network (Schoen et al. 1997).

The aim of this article is to investigate how family orientations shape the effect of fertility on SWB. Because attitudes and preferences impact gains in life satisfaction from having a child, we study whether family orientations work as a moderator of the fertility–SWB relationship. By doing so, we can also take into account that effects of fertility on SWB differ between people and are likely to be partially foreseen, so that those who have a child tend to be those who experience the most positive impact of childbearing. By using a propensity score matching approach, we aim to investigate the relationship between parenthood and SWB, taking into account potential sources of selection bias or endogeneity.

Our focus is on interindividual differences in SWB resulting from the birth of a child. Specifically, we aim to answer the following research question: *Is an individual who has a child more satisfied than his/her counterpart who does not have the child but has similar socioeconomic characteristics and family orientations?* Because childbearing effects differ significantly by parity, implying different expectations and costs (Billari et al. 2009), we consider two transitions: transition to first birth and transition to second birth. Thus, we compare people who become parents for the first time with those who stay childless, and people who have two children with those who have only one child.

Background and Hypotheses

Research has shown mixed results on the association between SWB and fertility, finding evidence for both positive and negative consequences of having children. Studies have also highlighted heterogeneous effects of fertility on SWB. Some studies have shown that the fertility–SWB relationship differs across cultural and institutional contexts (Aassve et al. 2012; Margolis and Myrskylä 2011); others have instead focused on individual-level characteristics, investigating the role of socioeconomic factors, such as age, gender, or economic well-being (Nelson et al. 2012). Although these studies provide important insights into the relationship between SWB and fertility, almost all of them have used a cross-sectional approach. They therefore suffer from the potential bias caused by endogeneity of childbearing behavior.

To overcome this drawback, Kohler et al. (2005) conducted a twin study,¹ finding that becoming a parent contributes positively to parents' happiness. Specifically, they found a nonlinear effect of children on happiness, especially for women. Women's happiness increases after the first child, but having additional children is not associated with further increases in well-being.

Recent studies have used longitudinal models with individual fixed effects to control for time-invariant unobserved variables. The longitudinal dimension of these studies has also been used to empirically test the set point theory, also referred to as the dynamic equilibrium theory (Headley and Wearing 1989). According to this theory, every individual has his/her own baseline level of SWB that is primarily determined by genetic factors and therefore cannot be permanently modified by life events, such as having a child. Clark et al. (2008), for example, showed that the birth of a child only temporarily changes an individual's level of SWB, with anticipation effects before the event and adaptation effects after childbearing, but no enduring long-term effects. Other studies have also found that SWB increases in the years surrounding childbearing and decreases thereafter (Keizer et al. 2010; Myrskylä and Margolis 2014; Pollmann-Shult 2014; Pouwels 2011). Moreover, some of the aforementioned studies examined how specific socioeconomic characteristics (e.g., age, gender, and education) may shape and moderate the effect of having a child on changes in SWB over time (Myrskylä and Margolis 2014), also focusing on the costs arising from different family contexts (Pollmann-Shult 2014).

In this study, we use an alternative method—propensity score matching—that also allows us to approximate the causal effect of having a child on SWB. Building on a study by Kravdal (2014), we take into account that heterogeneous family orientations may lead to heterogeneous effects of fertility on SWB. People with different family orientations have different values, gender attitudes, and preferences toward career and family; in turn, they have different expectations about how childbearing can affect their SWB.

According to the prospect theory (Kahneman and Tversky 1979), individuals who decide between uncertain alternatives, like having or not having a child, base their choice on the potential gain or loss they expect to receive when the event occurs. Thus, the probability of an individual's having a child also depends on the expectations that individual has about the impact parenthood has on his or her SWB. As Kravdal (2014)

¹ See Amin et al. (2015) for a discussion on the monozygotic twins fixed-effects estimators.

pointed out, this leads to the problem that the factors shaping the effect of the treatment “having a child” also affect the probability that the treatment takes place at all. That is, those who expect a higher gain in SWB from the birth of a child will be those who are more likely to actually have a child. To address this issue, we match individuals not only on several socioeconomic characteristics but also on family orientations.

By *family orientations* we mean two aspects that we believe are the main value-related driving forces influencing both the effect of having a child on an individual’s SWB and the actual fertility behavior of that individual: gender role attitudes and preferences toward work and family. As Hakim (2000) showed, these two factors have a significant influence on an individual’s fertility behavior. We envision them playing a relevant role also in shaping how the birth of a child affects an individual’s SWB: individuals with different family orientations have different expectations about how a child may influence their own SWB, and expectations may be differently met. Specifically, we divide individuals into three groups based on their family orientations: Traditional, Mixed, and Modern. This categorization is inspired by Hakim’s (2000, 2003) theoretical work and the empirical strategy adopted in several studies (e.g., Kan 2007; Sweeting et al. 2014).

The Traditional group is made up of men and women who endorse a male-breadwinner model, in which the female partner is expected to stay at home, not work, and be the main, if not the only, partner responsible for housework and childcare activities. The opposite group is Modern, which consists of people with a dual-earner family system and who follow a gender-egalitarian model. People with these family orientations think that both partners should work and contribute equally to household income, housework, and childcare.

Between the Traditional and Modern groups is a third group that, in advanced societies, is occupied by the majority of the population. This group follows an intermediate or mixed family model in which the female partner is not simply a housewife but likely works (at least part-time) in addition to being the main partner responsible for household and childcare activities.

By comparing those who have a child with their counterparts who do not within each of the three groups and for men and women separately, we address another problematic issue that Kravdal (2014) raised: looking at the average effect of having a child on SWB might be misleading because if differences in attitudes, preferences, and expectations are substantial, the average effect may conceal variations and become meaningless.

In the literature, we find different indicators, definitions, and metrics of SWB, each of which has its own specificities and foci (Diener et al. 1999; Dolan et al. 2007). We focus on life satisfaction as a measure of SWB because, as some scholars have noted (Dolan et al. 2007; Haller and Hadler 2006), life satisfaction is a multidimensional construct that focuses primarily on the cognitive aspects of SWB. We therefore believe that a life satisfaction measure may be particularly well suited to capturing an individual’s evaluation of the costs and benefits associated with childbearing in the long term.

Previous studies (e.g., Aassve et al. 2012; Nelson et al. 2012) have found that fathers more often have a higher level of SWB than mothers. We are interested in testing whether results from our matching framework also find support for this association. Men and women experience family and work life

differently, and women usually face greater work–family conflict and experience higher costs and stresses associated with childrearing, which may at least partially offset gains in life satisfaction. Therefore, we envision that *the gain in life satisfaction for fathers compared with nonfathers is larger than that for mothers compared with nonmothers* (Hypothesis 1).

Further, we think that differences across the three family orientations categories that we have identified are more relevant for women than for men. Men may have very different gender role attitudes and therefore expect their female partner to have a different role within and outside the household (e.g., housewife, working mother). They might even be more or less willing to share household and childcare activities with the partner. However, unlike women, they likely do not differ much across the three groups in terms of preferences toward work or family. Traditional, Mixed, and Modern men do not have significantly different lifestyles because they are expected to work anyway. As a consequence of persistent cultural gender differences, men usually do not have the option of choosing between family and work. Even if their participation in the childcare activities might be a result of either a conscious choice or a pressure from a working wife, we can assume that their sources of life satisfaction are likely more diversified than those of women because men's daily life is spent mainly outside the household. Thus, we propose the following hypothesis: *Family orientations play a stronger role among women than among men* (Hypothesis 2).

For women, we do expect great differences across the three groups of family orientations. Specifically, we believe that the number and type of sources of life satisfaction vary significantly across categories. The main assumption we make is that the relative effect of the benefits coming from motherhood on an individual's SWB varies according to the number and importance of the different sources of an individual's life satisfaction. While we expect Traditional women to place a very high value on family life and children as the main, if not the only, source of life satisfaction, we envision Modern women to have other sources of life satisfaction stemming from the work or career sphere, which may be even more important than having children. Thus, the gain in life satisfaction experienced from having a child should be higher for Traditional women than for Modern women. We believe that the sources of life satisfaction are more diversified for Mixed women than for Traditional women, likely because women from the Mixed group also have an active working life outside the household. However, they may still place a high value on family life, considering having children as still the primary source of life satisfaction. If this is true, this Mixed group may bear the highest costs of having children, derived from trying to reconcile and balance work and family. We therefore expect *mothers to be more satisfied than nonmothers among women who place a higher value on a life devoted to family and children: first Traditional, then Mixed, and finally Modern* (Hypothesis 3).

We also explore whether the effect of fertility on SWB changes by parity. Specifically, we estimate the effect of having a first child (versus being childless) and having the second child (versus having only one child). As Billari et al. (2009) argued, we assume that childbearing effects differ greatly by parity and imply different expectations and costs.

Data and Measurements

Data

We use the British Household Panel Survey (BHPS), an annual panel survey with a nationally representative sample of about 5,500 households recruited in 1991, containing approximately 10,000 interviewed individuals. Participants are reinterviewed each successive year for 18 years; participants who split from original households to form new ones are followed, and all adult members of these households are also interviewed. Similarly, new members joining sample households become eligible for interview, and children are interviewed beginning at age 16. The BHPS data set is well suited to investigating the relationship between fertility and life satisfaction because it provides information on several socioeconomic characteristics, family orientations, fertility history, and SWB measured over time.

Sample

To investigate the effect of having a first birth, we selected women and men aged 16 years or older who remained childless or had one child during the observation period. We also included in our sample individuals who had more than one child during the observation period, but we censored them two years before the birth of the second child to avoid any potential anticipation effects of the second child on SWB. We opted for this approach after checking that anticipation effects appear no sooner than the year before childbearing. After deletion of missing values, our working sample consists of 5,278 women and 5,127 men, corresponding to 21,762 and 19,600 person-year observations, respectively.

We used a similar strategy to identify the sample for the analysis focusing on the comparison between those who have two children and those who have only one. The final working sample consists of 1,505 women and 2,054 men, corresponding to 4,173 and 6,535 person-year observations, respectively.

Measurements

Dependent Variable

In the BHPS, life satisfaction is measured using the following question: “How dissatisfied or satisfied are you with your life overall,” with answers ranging from 1 (not satisfied at all) to 7 (completely satisfied). Unfortunately, this question was asked only from Wave 6 until the end of the survey (i.e., Wave 18), but with a gap in Wave 11. Therefore, we could not use Waves 1–5 and Wave 11 to measure our outcome variable.

Family Orientations

We measured family orientations using the following six items on 5-point scales (ranging from strongly agree to strongly disagree):

1. A preschool child is likely to suffer if his or her mother works.

2. All in all, family life suffers when the woman has a full-time job.
3. A woman and her family would all be happier if she goes out to work.
4. Both the husband and wife should contribute to the household income.
5. Having a full-time job is the best way for a woman to be an independent person.
6. A husband's job is to earn money; a wife's job is to look after the home and family.

These items, which refer to both gender role attitudes and preference toward work and family, are measured in alternating waves in the BHPS. Therefore, we imputed missing values using the average between the values observed in the two contiguous waves.

To construct a valid and reliable index measuring individuals' family orientations, we followed the strategy adopted by Kan (2007), who used the same items in the BHPS data set to evaluate Hakim's (2000) preference theory, and by Sweeting et al. (2014), who also used the same items and data to measure gender role attitudes. Specifically, after recoding all the items in the same direction (i.e., a low score in the answers indicates a traditional orientation), we summed them to obtain a score index, which was found to be reliable (Cronbach's alpha = .64).

Respondents who scored in the lower quartile were defined as Traditional; those in the upper quartile were defined as Modern; and those in between were identified as Mixed. This partition of the index leaving the Mixed group as the biggest is in line with Hakim's theory and is based on our assumption that the majority of the population should have a mixed family orientation.

Methods

Defining the Causal Effect of Interest

We used the potential outcome framework (Rubin 1974) to formalize the definition of the causal estimand of interest. The following discussion refers to the estimation of the effect of the first child on SWB, but it can be easily adapted to the analysis of the effect of the second child.

Our treatment variable is a binary indicator, D , that takes value 1 for individuals who had a first child within the observation period and 0 for childless people. SWB, our outcome of interest, measured at time t on the individual i is indicated as Y_{it} .

Under the potential outcome framework, each individual has two potential outcomes: $Y(1)_{it}$ indicates the SWB that would have resulted if the individual had the first child and $Y(0)_{it}$ indicates the outcome if he/she had stayed childless. For each individual, the causal effect of having the first child on SWB is defined as $Y(1)_{it} - Y(0)_{it}$. Because each individual is observed only in either the treatment or the control group, either $Y(1)_{it}$ or $Y(0)_{it}$ is observed for each individual, and the other quantity (the counterfactual SWB) must be estimated. The parameter (causal estimand) we aimed to estimate is the so-called average treatment effect on the treated (ATT) (Imbens 2004):

$$ATT = \left[Y(1)_{it} - Y(0)_{it} \middle| D_{i0} = 1 \right],$$

where $t = 0$ indicates the year of childbirth. The ATT can be interpreted as the estimated average difference between the observed level of life satisfaction for treated individuals (those who gave birth to the first child) and the level of life satisfaction they would have had if they had remained childless. We focus on the ATT because we think it is conceptually more interesting than alternative parameters. The ATT measures the average effect of having a first child for those who had a first child (compared with the counterfactual situation of not having a child). By contrast, the average treatment effect (ATE) can be thought of as an average of the ATT and ATU (average treatment effect on untreated) with weights equal to the percentage of treated and control individuals, respectively. Therefore, the ATE would estimate the effect of having the first child also for control individuals—that is, individuals who remained childless during the observation period. In our opinion, this makes the ATE less attractive because it would include in the population of interest individuals with no chance, or a very low chance, of having a child.

The Matching Procedure

For treated individuals, the counterfactual outcome, $Y(0)_{it}$, must be estimated. For this purpose, we used a matching approach: for each individual who had a child (in the treatment group), we looked for one or more individuals who did not have a child (in the control group) but who was the most similar to the individual in the treatment group in terms of observed characteristics. Propensity score matching (PSM; Rosenbaum and Rubin 1983) is used to generate such a control group for each treated person. This technique has two steps. First, a model is estimated for the probability of experiencing the treatment (having a first child) as a function of confounding variables. Predicted probabilities (i.e., estimates of the propensity scores) are used to match individuals. In fact, Rosenbaum and Rubin (1983) showed that individuals with similar values of the propensity score have, in expectation, the same distribution of all of the covariates entering the propensity score. This *balancing property* makes it possible to compare units on the basis of the univariate propensity score instead of the multivariate set of covariates. In the second step, after the matching has been implemented, the outcomes of treated and matched control units are compared and causal effects are estimated using a simple weighted difference of means. Given that this second step is nonparametric, the method is considered to be semiparametric.

Matching can be implemented in several ways (Caliendo and Kopeinig 2008). As suggested in the treatment effects literature, we experimented with different matching algorithms and different specifications of the propensity score. The method we selected is radius matching with a caliper equal to one-quarter of the standard deviation of the estimated propensity score, because it guarantees the best balancing of covariates.

We estimated propensity scores using a logit model with several independent variables that are relevant for the mechanism that leads to childbearing and for the outcome under study (life satisfaction). Specifically, we included age, gender, marital status, health status and health problems, household income, education, employment status, and all six items measuring individuals' family orientations, as previously defined. Instead of matching on the three identified groups of family orientations

(i.e., Traditional, Mixed, and Modern), we used all six items as separate covariates in order to have a finer match on family orientations.

Exploiting the longitudinal structure of the data, we developed our estimator in a pre-post treatment setting, aiming to match on covariates measured two years before the treatment (childbirth).² Figure 1 illustrates our general matching procedure (panel a) and provides an example (panel b). We start by observing that an individual had a child at a given point in time ($t = 0$). For each treated individual, the matching algorithm finds one or more similar individuals in the control group in terms of propensity scores with covariates measured two years before childbirth ($t = -2$). In other words, the treated individual is matched with one or more control individuals at time $t = -2$. Then, outcomes of these individuals are compared at different points in time starting from $t = -1$ (one year before childbirth for the treated individual, but simply one year after the matching year for control individuals).³ We follow along this line by comparing outcomes of treated and control individuals until Wave 18 (or until they remain in the survey). To avoid contamination effects, we retain control units in the matched data set until two years before they (eventually) have a child.

Panel b of Fig. 1 offers an illustrative example. Consider an individual who gave birth at Wave 9. We match the treated individual with one (or more) similar control individual(s) using covariates measured at Wave 7. If these individuals do not drop out of the survey (and the control individuals remain childless) we can measure their SWB until Wave 18 (i.e., when the child of the treated individual is 9 years old). In this case, we cannot estimate the effect of childbirth on SWB at $t = +2$ because, as mentioned earlier, life satisfaction was not measured at Wave 11.

In principle, we could calculate ATTs comparing the SWB of treated and control individuals in each year starting from the year before childbirth (if both treated and control units are still in the survey). However, because the number of treated units considerably decreases over time (see Table 7 in the appendix), we investigated potential gains or losses in life satisfaction after childbirth not year by year, but measuring the ATT over longer periods. Specifically, we focused on life satisfaction measured over the entire period of observation (from the year before childbearing until the end of the period of observation); life satisfaction only before the treatment (i.e., the year before childbearing); and life satisfaction since the birth of the child. Furthermore, we aimed to study whether life satisfaction may vary according to different childrearing phases. To do so, we segmented the period over which we observe life satisfaction after the birth of a child, according to the child's age. We specifically looked at life satisfaction when the child was between 0 and 3 years old and between 4 and 11 years old. Thus, for example, for the period when the child was 4–11 years old, we averaged all the ATTs estimated comparing SWB measured at $t = 4, 5, 6, \dots, 11$ instead of presenting them separately at each time point.

As in all longitudinal studies, in our matching design, we observe different individuals for a different number of waves (an individual is followed for an average of 8.6 and 7.8 years in the samples for the first and second child, respectively). This implies

² In previous analyses, we also tried to match individuals three, four, and five years before the treatment, but we were able to observe diverging patterns of life satisfaction resulting from anticipation effects only in the year before the treatment. Thus, we opted to match two years before childbearing in order to follow individuals and their life satisfaction for a longer period after childbearing.

³ In principle, t can assume values in $\{-1, 0, 1, \dots, 11\}$.

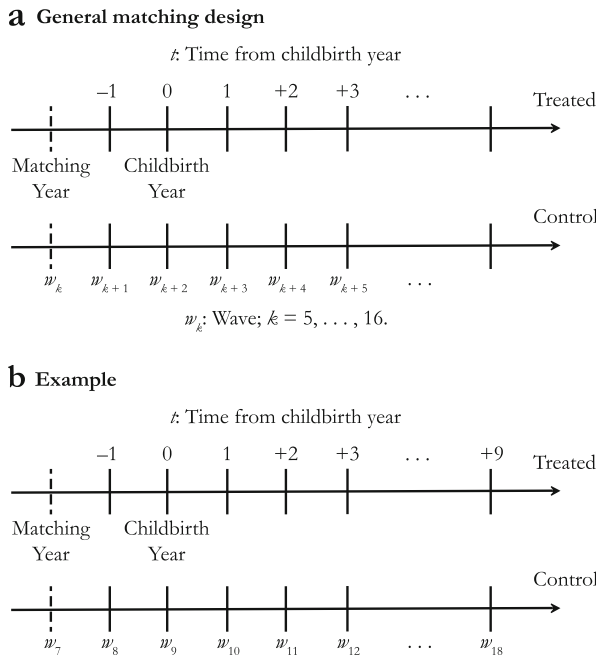


Fig. 1 Time lines describing the matching procedure for both the general design and an example. In the general matching design (panel a), we look for comparable control individuals for each treated person, with matching occurring two years before childbirth. In the example (panel b), childbirth occurs at Wave 9 and matching is at Wave 7. The outcome is measured from Wave 8 (in the year before childbirth) to Wave 18 (when the child is 9 years old). The outcome is missing at Wave 11 because the BHPS did not include the question on life satisfaction during that wave

that we use different samples when we estimate ATTs over different periods. In principle, this is not problematic if the number of waves each individual stays in the survey is not correlated with fertility behaviors and SWB. Indirect evidence supports this assumption. We compared the distribution of covariates across the different samples used in the analyses that refer to different periods (entire period, before childbearing, and so on) and found that the different samples had very similar average values for each covariate considered. Moreover, for control individuals, we found no pattern in life satisfaction over time. In other words, the number of waves for which control individuals are observed in the survey and their SWB are not correlated (see Tables 1 and 2, and Table 7 in the appendix).

Strengths and Limitations of the PSM Implemented on Longitudinal Data

As with all methods, PSM has advantages and limitations. PSM is more robust than parametric regression models, especially when the distributions of covariates in the two groups being compared are very different because the regression estimator depends heavily on extrapolation using the specific functional form of the model (Imbens 2014). In addition, PSM can highlight initial differences in observed covariates among treated and control units and offers the opportunity to check whether it was successful in reducing these imbalances (see the next section). Moreover, implementing matching on

Table 1 Life satisfaction and sample sizes for the analysis of the effect of the first child, over different periods of time

Time	Control			Treated		
	Number of Observations	Mean	SD	Number of Observations	Mean	SD
Entire Period	30,597	5.14	0.97	1,612	5.40	0.92
Before Childbirth	24,136	5.17	1.17	1,203	5.48	1.12
After Childbirth	25,416	5.13	0.97	1,556	5.37	0.98
Child 0–3	25,259	5.13	1.02	1,555	5.37	1.00
Child 4–11	12,335	5.09	0.97	365	5.16	1.00

longitudinal data, as we did, offers several advantages (Arpino and Aassve 2013). First, we can match on covariates measured before childbirth (two years before).⁴ In this way, we avoid controlling for covariates that could be on the causal pathway between the treatment and the outcome (mediators), generating biased estimates (e.g., Imbens 2004; Rosenbaum 1984). Second, the lagged value of the outcome variable can be included in the set of matching covariates. Matching on the lagged outcome, similar to longitudinal models with individual fixed-effects, allows us to control for time-invariant unobserved confounders, such as genetic factors or personal traits that could be correlated with life satisfaction (for detailed methodological discussion on this and related estimators, see Arpino and Aassve 2013; Athey and Imbens 2006; Imai and Kim 2014). The underlying assumption is that individuals in the treatment and control group can have different levels of SWB before childbearing. Comparing their SWB after childbearing would give a biased estimate because of these pretreatment differences. Matching on the pretreatment value of the outcome (SWB) eliminates this bias under the assumption that the pretreatment difference in SWB is due only to unobserved time-invariant factors (apart from *observable* time-variant factors).⁵ For practical reasons, instead of matching on the same question used for the outcome definition (life satisfaction), we matched on happiness.⁶ As noted earlier, life satisfaction was measured in the BHPS only starting from the sixth wave, so matching on life satisfaction would have implied losing one entire wave of observations, with a conspicuous decrease in the sample size of the treatment groups. Instead, we used happiness as a covariate and matched individuals starting from Wave 5 (first possible childbirth year is Wave 7). Happiness is a good substitute for life satisfaction because the two measures are highly correlated and offer consistent results (see Myrskylä and Margolis 2014).

Despite these important advantages of PSM implemented on longitudinal data, we acknowledge some limitations of the method. First, in the presence of unobserved *time-*

⁴ The average lag between the start of the observation period and the year of matching is 2.8 and 1.1, respectively, for the analyses on the first and second child. The average number of years of follow-up after matching for the two analyses is 3.8 and 5.1, respectively.

⁵ This assumption, which is a weaker version of the so-called unconfoundedness or conditional independence assumption, can be written for identification of ATT in our case as: $E[Y(0)_i | D_i = 1, X_{-2}, Y_{-2}] = E[Y(0)_i | D_i = 0, X_{-2}, Y_{-2}]$ (see, e.g., Arpino and Aassve 2013).

⁶ The question about happiness is as follows: “Have you recently been feeling reasonably happy, all things considered?” Responses range from 1 (much less happy than usual) to 4 (more happy than usual).

Table 2 Life satisfaction and sample sizes for the analysis of the effect of the second child, over different periods of time

Time	Control			Treated		
	Number of Observations	Mean	SD	Number of Observations	Mean	SD
Entire Period	7,482	4.98	1.07	1,481	5.27	0.88
Before Childbirth	5,898	5.02	1.26	1,059	5.37	1.10
After Childbirth	6,317	4.96	1.08	1,441	5.26	0.91
Child 0–3 years	6,311	4.98	1.11	1,439	5.27	0.94
Child 4–11 years	3,412	4.93	1.13	736	5.14	0.91

variant confounders, the PSM estimator is biased. We tried to overcome this issue by matching on a wide set of covariates that previous studies considered important predictors of childbearing and SWB. However, we cannot completely rule out the possibility that unobserved confounders exist. Second, if some covariates are endogenous, this would bias the estimates. Our measurement of covariates two years before childbirth reduces the extent of this problem.

Results

Descriptive Statistics and Covariates Balance

Tables 1 and 2 show the described time periods we focus on in our analysis. They also report the average level of life satisfaction and relative sample size of both the treated and the control groups.

As we mentioned earlier, because the aim of our study was to uncover whether gender and family orientations play any role in shaping the effect of childbearing on life satisfaction, we ran separate analyses by gender. Then, for women and men separately, we also ran different analyses for the three family orientations groups. We performed the matching within each subgroup, each time checking that the balancing of covariates was satisfactory.

As a measure of balance for the variables used for matching, we used the percent standardized bias, which for each covariate is defined as the standardized difference between the means of that variable in the treatment and the control groups (Rosenbaum and Rubin 1985). Values above 10 % are considered unacceptable (Normand et al. 2001). Table 8 in the appendix reports a general description of our sample and shows the balance of independent variables across treated and control units. Specifically, for each variable, we report the mean in the treatment and control groups and their standardized difference before and after matching. This table refers to the first analysis we conducted, in which we estimated the effect of the first child on the entire sample. The balance of independent variables for the other analyses is similar but not reported, for the sake of brevity. Table 8 shows that, before matching, substantial imbalance existed between covariates. Matching was highly

satisfactory, given that balance improved considerably. After matching, none of the covariates showed a bias higher than 4 %.

PSM Estimates

Tables 3, 4, 5 and 6 report ATT estimates from PSM along with confidence intervals for pairwise comparisons. These intervals are centered on the ATT estimate and have lengths equal to $2 \times 1.39 \times$ standard error. As Goldstein and Healy (1995) showed, this is necessary in order to have an average level of 5 % for the type I error probability in the pairwise comparisons of a group of means. Therefore, an overlap in terms of confidence intervals indicates that ATTs are not significantly different, whereas nonoverlap reflects that they are.

Tables 3 and 4 report results concerning the effect of the first child—that is, those in which the treatment is the birth of the first child, and therefore the comparison is between parents of one child and childless individuals. As shown in Table 3, looking at the entire sample, parents seem to be more satisfied than nonparents, but the gain in life satisfaction derived from parenthood seems to be short-lived. Parents seem to have higher life satisfaction than childless individuals until the child is 3 years old; thereafter, no positive effect is evident. Previous research found similar results (Clark et al. 2008). In line with our first hypothesis, gains in life satisfaction for fathers versus nonfathers seem to be higher than those for mothers versus nonmothers. Confidence

Table 3 Average treatment effect among the treated (ATT) on life satisfaction from propensity score matching considering as treatment the birth of the first child: Analysis on the entire sample and by gender

Time		Total Sample	Women	Men
Entire Period	ATT	0.21***	0.12**	0.26***
	CI	(0.17, 0.24)	(0.07, 0.17)	(0.21, 0.31)
	<i>N</i>	32,209	17,005	15,202
Before Childbearing	ATT	0.25***	0.16**	0.31***
	CI	(0.20, 0.30)	(0.09, 0.23)	(0.25, 0.38)
	<i>N</i>	25,339	13,232	12,105
After Childbearing	ATT	0.19***	0.09*	0.25***
	CI	(0.15, 0.23)	(0.04, 0.15)	(0.20, 0.30)
	<i>N</i>	26,972	14,298	12,673
When Child Is 0–3 Years	ATT	0.20***	0.10*	0.25***
	CI	(0.16, 0.23)	(0.04, 0.16)	(0.19, 0.30)
	<i>N</i>	26,814	14,193	12,620
When Child Is 4–11 Years	ATT	0.04	–0.05	0.10
	CI	(–0.04, 0.11)	(–0.06, 0.16)	(0.00, 0.21)
	<i>N</i>	12,700	6,829	5,871

Notes: *N* is the sample size of each subgroup; CI is the confidence interval for 5 % level pairwise comparisons of ATTs.

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 4 Average treatment effect among the treated (ATT) on life satisfaction from propensity score matching considering as treatment the birth of the first child: Analysis on women and men separately, by family orientations

Time	Women			Men		
	Traditional	Mixed	Modern	Traditional	Mixed	Modern
Entire Period	ATT	0.09	0.07	0.29***	0.24***	0.32***
	CI	(0.23, 0.47)	(0.02, 0.15)	(-0.05, 0.19)	(0.17, 0.41)	(0.19, 0.45)
	N	2,743	11,449	2,766	2,542	2,375
Before Childbearing	ATT	0.28*	0.14*	0.12	0.21*	0.43***
	CI	(0.12, 0.44)	(0.06, 0.22)	(-0.07, 0.30)	(0.08, 0.34)	(0.26, 0.60)
	N	2,121	8,958	2,120	2,029	1,879
After Childbearing	ATT	0.37***	0.06	0.02	0.28**	0.23*
	CI	(0.24, 0.50)	(0.00, 0.13)	(-0.11, 0.15)	(0.14, 0.42)	(0.09, 0.37)
	N	2,338	9,652	2,271	2,105	1,966
When Child Is 0–3 Years	ATT	0.41***	0.07	0.02	0.27**	0.28**
	CI	(0.28, 0.54)	(0.00, 0.14)	(-0.11, 0.15)	(0.14, 0.40)	(0.14, 0.42)
	N	2,325	9,578	2,253	2,090	1,956
When Child Is 4–11 Years	ATT	0.07	-0.06	0.03	0.06	0.24
	CI	(-0.32, 0.45)	(-0.19, 0.07)	(-0.19, 0.25)	(-0.19, 0.32)	(-0.07, 0.55)
	N	1,132	4,654	1,035	999	870

Notes: N is the sample size of each subgroup; CI is the confidence interval for 5 % level pairwise comparisons of ATTs.

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 5 Average treatment effect among the treated (ATT) on life satisfaction from propensity score matching considering as treatment the birth of the second child: Analysis on the entire sample and by gender

Time		Total Sample	Women	Men
Entire Period	ATT	0.16***	0.07	0.23***
	CI	(0.12, 0.20)	(0.01, 0.13)	(0.18, 0.39)
	<i>N</i>	8,693	3,438	5,222
Before Childbearing	ATT	0.20***	0.10	0.28***
	CI	(0.15, 0.26)	(0.02, 0.18)	(0.20, 0.35)
	<i>N</i>	6,957	2,644	4,312
After Childbearing	ATT	0.17***	0.07	0.24***
	CI	(0.13, 0.21)	(0.01, 0.13)	(0.18, 0.30)
	<i>N</i>	7,758	2,945	4,811
When Child Is 0–3 Years	ATT	0.16***	0.08	0.21***
	CI	(0.12, 0.21)	(0.01, 0.14)	(0.15, 0.27)
	<i>N</i>	7,750	2,940	4,808
When Child Is 4–11 Years	ATT	0.10*	0.00	0.18**
	CI	(0.04, 0.16)	(-0.09, 0.09)	(0.10, 0.26)
	<i>N</i>	4,148	1,538	2,608

Notes: *N* is the sample size of each subgroup; CI is the confidence interval for 5 % level pairwise comparisons of ATTs.

* $p < .05$; ** $p < .01$; *** $p < .001$

intervals suggest that this gender difference is significant. Mothers show a strong positive effect on life satisfaction the year before the birth of the first child. However, the gain in life satisfaction during the first three years after childbearing becomes very small, and the effect of parenthood on life satisfaction for mothers even becomes negative, although not significant, when the child is 4 years or older. By contrast, fathers continue to show significant gains in life satisfaction compared with nonfathers; as in the total sample, however, the positive effect becomes statistically nonsignificant when the child is 4–11 years old.

The next step is to uncover whether the average effects on life satisfaction for men and women conceal variations that somehow balance them out (Kravdal 2014). We therefore compare parents and nonparents within each of three family orientations groups. Table 4 shows the results for women and men by family orientations. We found that the gains in life satisfaction for the total sample of women is driven only by the high gains for Traditional mothers compared with their childless counterparts, in line with our hypothesis. Although Modern mothers do not show a positive effect either before or after childbearing, Mixed mothers show a significant and positive anticipation effect before having the first child but no effects after the treatment. Actually, the effect is negative albeit not significant when the child is 4 years old or older. These findings seem to support our reasoning that Mixed women bear higher costs of childbearing than women in the other family orientations groups. The

Table 6 Average treatment effect among the treated (ATT) on life satisfaction from propensity score matching considering as treatment the birth of the second child. Analysis only on women, by family orientations

Time	Women				Men			
	Traditional	Mixed	Modern	N	Traditional	Mixed	Modern	N
Entire Period	ATT	0.12	0.06	0.13	0.22*	0.24***	0.12	
	CI	(0.00, 0.25)	(-0.01, 0.13)	(-0.04, 0.30)	(0.09, 0.34)	(0.18, 0.31)	(-0.02, 0.27)	
	N	589	2,186	603	844	3,525	1,147	
Before Childbearing	ATT	0.17	0.12	0.03	0.21	0.27***	0.28*	
	CI	(-0.02, 0.35)	(0.02, 0.23)	(-0.19, 0.24)	(0.05, 0.38)	(0.18, 0.37)	(0.10, 0.46)	
	N	438	1,687	476	659	2,735	916	
After Childbearing	ATT	0.07	0.07	0.17	0.32***	0.26***	0.05	
	CI	(-0.05, 0.19)	(-0.01, 0.14)	(-0.02, 0.35)	(0.18, 0.46)	(0.19, 0.33)	(-0.11, 0.21)	
	N	520	1,873	499	735	3,081	988	
When Child Is 0-3 Years	ATT	0.05	0.10	0.16	0.26*	0.25***	0.05	
	CI	(-0.08, 0.18)	(0.02, 0.17)	(-0.01, 0.34)	(0.11, 0.41)	(0.17, 0.32)	(-0.12, 0.21)	
	N	518	1,871	498	735	3,078	988	
When Child Is 4-11 Years	ATT	-0.06	-0.08	0.33	0.33**	0.17*	-0.07	
	CI	(-0.26, 0.13)	(-0.19, 0.03)	(0.02, 0.64)	(0.16, 0.50)	(0.08, 0.26)	(-0.29, 0.15)	
	N	282	995	233	430	1,670	504	

Notes: N is the sample size of each subgroup; CI is the confidence interval for 5 % level pairwise comparisons of ATTs.

* $p < .05$; ** $p < .01$; *** $p < .001$

finding that Mixed mothers have a rather high, positive anticipation effect but no gain in life satisfaction after childbirth may indeed result from high costs, such as time costs (Pollmann-Shult 2014) and difficulties in balancing work and family life, that offset the higher satisfaction derived from parenthood. That is, Mixed mothers might experience a great mismatch between their expectations and their actual outcomes in terms of life satisfaction after the birth of the first child. On the other hand, Modern women, who seem never to experience a gain in life satisfaction (not even before childbearing) may have other relevant sources of life satisfaction (for example, work and career) that overall reduce the relative role of parenthood for life satisfaction, even when expectations from childbearing are not low. Thus, the SWB of Modern mothers is not affected by the birth of a child because the factors in their “utility function” are various and change in only one of those (i.e., having a child) may not significantly affect their life satisfaction.

For men, we find no substantial differences in terms of effects on life satisfaction for the three groups under study. Fathers always seem more satisfied than nonfathers both before and after childbearing and regardless of their family orientations. This is in line with our second hypothesis. We believe that differences across categories of family orientations are less relevant for men than for women because the attitudinal aspect related to the preferences toward work or family is less applicable for men than for women. Men may have very different gender role attitudes but likely not very different lifestyles because they are expected to work anyway. As a consequence of persistent cultural gender differences, they usually do not have the socially acceptable option of choosing between staying at home and going to work.

Tables 5 and 6 report results regarding the effect of having a second child. These findings stem from the comparison between those who have two children and those who only have one. Table 5 shows that, looking at the entire sample, parents who have a second child seem to gain more in terms of life satisfaction than those who have only one child. The positive effect, different from the first child, seems to last in the long term, even when the second child is between 4 and 11 years old. It is very interesting to see that this positive effect observed in the total sample is entirely due to the gain in life satisfaction experienced by fathers. Mothers do not show any effect, regardless of their family orientations, as Table 6 reports. This finding is in line with results of previous research (Myrskylä and Margolis 2014), which shows that the second child has a much weaker effect on life satisfaction. Table 6 also shows that Traditional and Mixed fathers see a positive and enduring effect on life satisfaction due to the birth of their second child, whereas Modern fathers have a trend in life satisfaction that is similar to that of Mixed mothers with the birth of the first child. Modern fathers experience an increase in life satisfaction the year before childbearing, but then they do not have any gain after the birth of the child. We believe this may be due to the higher costs that Modern fathers experience because they are likely willing to share household and childcare tasks. As a consequence, because costs increase with the number of children, the arrival of the second child adds to their feelings of burden from reconciling work and family life and from time constraints.

Conclusions

The aim of our study was twofold. First, we attempted to extend the literature on fertility and SWB by studying how heterogeneous family orientations for men and women can differently shape the effect of having a child on an individual's life satisfaction. By doing so, we were able to take into account that effects of fertility on SWB differ between people and are likely to be partially foreseen; we were able to control for the fact that those who eventually have a child tend to be those for whom childbearing has the most positive impact.

Second, we used a propensity score matching (PSM) approach with longitudinal data to investigate how the birth of the first or second child impacts an individual's SWB. This method allowed us to approximate the causal effect of fertility on life satisfaction as the within-individual longitudinal approach does, by controlling for factors affecting both fertility and happiness. Using PSM allowed us to compare individuals who had different numbers of children but very similar socioeconomic characteristics and family orientations before the childbearing events.

Inspired by Hakim's (2000) preference theory, we categorized individuals into three groups based on their family orientations: Traditional, Mixed, and Modern. In comparing life satisfaction of individuals who had one child (or two children) with those who did not (or had only one child), we found significant differences between men and women and across different family orientations groups.

Overall, we found that individuals who had a first (or second) child were more satisfied than those who remained childless (or had one child only). Focusing on the first child, we found two interesting results. First, the gain in life satisfaction for fathers compared with nonfathers was bigger than that experienced by mothers compared with nonmothers. Second, only Traditional mothers seemed to be more satisfied than their childless counterparts. However, although Modern mothers never experienced a gain in life satisfaction, Mixed mothers seemed more satisfied than Mixed nonmothers the year before the birth of the first child but did not have a gain in life satisfaction after childbearing. We think this result is due to the high costs that Mixed mothers bear (given their likelihood of both working outside of home and being the main partner responsible for family care) with the arrival of a child. On the other hand, we did not find any relevant difference across the three family orientations groups among men: fathers always seemed more satisfied than nonfathers.

We found that fathers of a second child were more satisfied than fathers of only one child, whereas no gain in life satisfaction was observed for mothers of a second child compared with mothers of only one child only, regardless of their family orientations. Results showed that Traditional and Mixed fathers are more satisfied than their nonfather counterparts; however, Modern fathers, similar to Mixed mothers with the first child, experienced a gain in life satisfaction before the birth of the second child but not afterward. We believe this finding is due to the fact that Modern fathers are more willing to help with housework and childcare activities, leading them to perceive higher costs and stresses associated with the arrival of a second child.

This study allowed us to uncover whether the overall average effects of having a child on life satisfaction for men and women conceal underlying variations. Our findings are

important because they show that the effect of childbearing on SWB substantially changes according to an individual's family orientations, especially among women. This aspect should help address family policies aimed at improving the SWB of mothers. Particularly interesting is that the largest group of mothers—mothers with Mixed family orientations—is the one that does not experience an increase in SWB after the birth of the first child compared with their childless counterparts but does experience higher life satisfaction before childbearing. It seems that for this group of women, expectations about childbearing are hardly met, leading to a decrease in life satisfaction. Family policies that specifically target this group of women could help them meet expectations and thereby increase their life satisfaction. Moreover, the mismatch between the pre- and post-childbearing SWB that we found in women with Mixed family orientations suggests that potential parents may actually not be well informed about the consequences of childbearing, thereby undermining the assumption that individuals can foresee childbearing consequences. As Margolis and Myrskylä (2015) pointed out, such a mismatch reduces the probability of having a second child.

Our study suffers from four limitations. First, although our period of observation is rather long, at 13 years, we cannot exclude the possibility that those individuals who we consider childless or with only one child may have had a child after the window of observation. This would by no means bias our results, but we cannot have a full fertility history for all the individuals in our sample. Second, the anticipation effect that we observed before the birth of the child may be partially mixed up with a “pregnancy effect.” Some individuals may have been interviewed the year before childbearing when they were already pregnant. In these cases, the pure anticipation effect and the effect of being pregnant on an individual's SWB cannot be disentangled. Third, our categorization by family orientations does not fully solve the treatment-heterogeneity problem (i.e., the fact that the people who have a child are those who are more likely to be positively affected by childbearing). Each family orientations group contains further variations—perhaps especially for men, because the items used to identify the different family orientations are more suitable for women than for men. The orientation of the items toward women's experiences could partially explain why we did not find any variation in SWB among the three family orientations groups for men. Fourth, we were not able to distinguish between direct and indirect consequences of the treatment (i.e., having a child) on an individual's well-being. For instance, we cannot know whether the fact that we find parents more satisfied than nonparents is a direct consequence of the birth of the child or whether it is due to the fact that nonparents never found the right partner with whom to form a family. Moreover, it would have been interesting to also include actual behavioral information—for example, on the division of household tasks—but this was not possible because the individual was our unit of analysis. Behavioral measures, of course, were not available for the several individuals who were not yet in a stable relationship at the time of matching. We hope further research focuses on couples and includes, perhaps together with family orientations, behavioral measures of gender roles and lifestyle.

Our study contributes to the literature on the relationship between fertility and SWB in integrating the important role of family orientations and explaining why and how heterogeneity in gender attitudes and lifestyle preferences can lead to heterogeneous effects of having a child on an individual's life satisfaction. Moreover, we used PSM, which, to the best of our knowledge, has never been used to study this topic. We explained the relevant advantages of this method compared with those already used in the literature, and we hope

that further research on this topic will make use of this technique. Finally, by using PSM and by taking heterogeneous family orientations into account, we overcame a potential source of bias that Kravdal (2014) pointed out as characterizing the literature on the topic. Our approach can be used to control for the fact that individuals with different family orientations—and therefore different expectations about the effect of having a child on their own SWB—may be more or less willing to have one or more children.

Acknowledgments The authors would like to thank Letizia Mencarini and Arnstein Aassve for their useful comments. Moreover, the authors are very grateful to the anonymous reviewers and the editor for the careful review of the manuscript.

Compliance with Ethical Standards

Funding The authors gratefully acknowledge financial support from the European Research Council under the European ERC Grant Agreement no StG-313617 (SWELL-FER: Subjective Well-being and Fertility, P.I. Letizia Mencarini).

Appendix

Table 7 Life satisfaction and sample sizes year by year for the analysis of the effect of the first child on life satisfaction

Year	Control			Treated		
	Number of Observations	Mean	SD	Number of Observations	Mean	SD
-1	24,136	5.17	1.17	1,203	5.48	1.12
0	20,139	5.15	1.16	1,294	5.51	1.10
1	17,587	5.13	1.15	975	5.26	1.16
2	15,124	5.11	1.15	641	5.12	1.12
3	13,254	5.09	1.16	476	5.14	1.15
4	10,992	5.09	1.14	342	5.12	1.16
5	8,972	5.09	1.13	253	5.15	1.11
6	7,165	5.08	1.12	199	5.14	1.02
7	5,764	5.08	1.12	166	5.04	1.19
8	4,546	5.10	1.12	124	5.08	1.19
9	4,222	5.09	1.13	120	5.19	1.00
10	3,418	5.08	1.13	104	5.10	1.15
11	2,698	5.06	1.13	76	5.24	1.13

Table 8 Balance of independent variables before and after matching: Pooled sample for the analysis of the effect of the first child

Variable	Sample	Mean		% Bias	Variable	Sample	Mean		% Bias
		Treated	Control				Treated	Control	
Gender	Unmatched	0.57	0.47	19.20	Edu ISCED 3	Unmatched	0.41	0.49	-14.70
	Matched	0.57	0.57	-1.10		Matched	0.41	0.41	0.00
Age	Unmatched	27.33	25.40	30.20	Edu ISCED 4	Unmatched	0.10	0.09	5.20
	Matched	27.33	27.33	0.00		Matched	0.10	0.11	-2.10
Age Squared	Unmatched	778.91	695.01	23.60	Edu ISCED 5	Unmatched	0.02	0.04	-10.80
	Matched	778.91	778.37	0.20		Matched	0.02	0.02	0.00
Happiness	Unmatched	3.06	3.06	0.70	Employed	Unmatched	0.87	0.65	53.90
	Matched	3.06	3.06	0.40		Matched	0.87	0.87	0.10
Single	Unmatched	0.27	0.71	-99.20	Unemployed	Unmatched	0.06	0.06	-2.80
	Matched	0.27	0.28	-1.80		Matched	0.06	0.06	-0.70
Married	Unmatched	0.45	0.11	80.30	Other LFS	Unmatched	0.07	0.29	-58.60
	Matched	0.45	0.45	-0.10		Matched	0.07	0.07	0.30
Cohabiting	Unmatched	0.27	0.16	28.10	Item 1 Attitudes	Unmatched	2.80	2.90	-10.90
	Matched	0.27	0.26	2.20		Matched	2.80	2.79	0.70
Separated	Unmatched	0.01	0.02	-3.10	Item 2 Attitudes	Unmatched	3.42	3.48	-6.30
	Matched	0.01	0.01	-0.50		Matched	3.42	3.42	0.40
Health Status	Unmatched	1.99	1.99	0.70	Item 3 Attitudes	Unmatched	3.07	3.05	3.60
	Matched	1.99	2.01	-2.10		Matched	3.07	3.06	1.00
Health Problems	Unmatched	0.45	0.41	7.60	Item 4 Attitudes	Unmatched	2.40	2.32	10.10
	Matched	0.45	0.46	-2.00		Matched	2.40	2.38	2.80

Table 8 (continued)

Variable	Sample	Mean		% Bias	Variable	Sample	Mean		% Bias
		Treated	Control				Treated	Control	
Income	Unmatched	2,832.00	3,012.60	-9.60	Item 5 Attitudes	Unmatched	2.78	2.74	4.30
	Matched	2,832.00	2,819.10	0.70		Matched	2.78	2.77	1.00
Edu ISCED 1	Unmatched	0.21	0.17	10.90	Item 6 Attitudes	Unmatched	3.98	3.99	-1.30
	Matched	0.21	0.22	-2.60		Matched	3.98	3.98	-0.10
Edu ISCED 2	Unmatched	0.25	0.21	8.00					
	Matched	0.25	0.23	3.90					

Note: The table reports the mean of each covariate in the treated and control group (third and fourth columns, respectively) and the percent standardized difference between the two (fifth column) before and after matching.

References

- Aassve, A., Goisis, A., & Sironi, M. (2012). Happiness and childbearing across Europe. *Social Indicator Research, 108*, 65–86.
- Amin, V., Behrman, J. R., Kohler, H. P., Xiong, Y., & Zhang, J. (2015). Causal inferences: Identical twins help and clarity about necessary assumptions is critical. *Social Science & Medicine, 127*, 201–202.
- Arpino, B., & Aassve, A. (2013). Estimating the causal effect of fertility on economic wellbeing: Data requirements, identifying assumptions and estimation methods. *Empirical Economics, 44*, 355–385.
- Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica, 74*, 431–497.
- Billari, F. C. (2009). The happiness commonality: Fertility decision in low-fertility settings. In UNECE (Ed.), *How generations and gender shape demographic change* (pp. 7–38). New York, NY, and Geneva, Switzerland: United Nations.
- Billari, F. C., Philipov, D., & Testa, M. (2009). Attitudes, norms and perceived behavioural control: Explaining fertility intentions in Bulgaria. *European Journal of Population, 25*, 439–465.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*, 31–72.
- Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *Economic Journal, 118*, F222–F243.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302.
- Dolan, P., Peasgood, T., & White, M. (2007). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology, 29*, 94–122.
- Friedman, D., Hechter, M., & Kanazawa, S. (1994). A theory of the value of children. *Demography, 31*, 375–401.
- Goldstein, H., & Healy, M. J. R. (1995). The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A, 158*, 175–177.
- Hakim, C. (2000). *Work-lifestyle choices in the 21st century*. Oxford, UK: Oxford University Press.
- Hakim, C. (2003). A new approach to explaining fertility patterns: Preference theory. *Population and Development Review, 29*, 349–374.
- Haller, M., & Hadler, M. (2006). How social relations and structures can produce happiness and unhappiness: An international comparative analysis. *Social Indicators Research, 75*, 169–216.
- Hoffman, L. W., & Hoffman, M. L. (1973). The value of children to parents. In J. T. Fawcett (Ed.), *Psychological perspectives on population* (pp. 19–76). New York, NY: Basic Books.
- Hoffman, L. W., & Manis, J. D. (1979). The value of children in the United States: A new approach to the study of fertility. *Journal of Marriage and the Family, 41*, 583–596.
- Imai, K., & Kim, I. S. (2014). On the use of linear fixed effects regression estimators for causal inference. Unpublished manuscript, Department of Politics, Princeton University, Princeton, NJ. Retrieved from <http://imai.princeton.edu/research/files/FEmatch.pdf>
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics, 86*, 4–29.
- Imbens, G. W. (2014). *Matching methods in practice: Three examples* (NBER Working Paper No. 19959). Cambridge, MA: National Bureau of Economic Research.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
- Kan, M. Y. (2007). Work orientation and wives' employment careers: An evaluation of Hakim's preference theory. *Work and Occupations, 34*, 430–462.
- Keizer, R., Dykstra, P. A., & Poortman, A. R. (2010). The transition to parenthood and well-being: The impact of partner status and work hour transitions. *Journal of Family Psychology, 4*, 429–438.
- Kohler, H.-P., Behrman, J. R., & Skytthe, A. (2005). Partner + children = happiness? The effects of partnerships and fertility on well-being. *Population and Development Review, 31*, 407–445.
- Kravdal, Ø. (2014). The estimation of fertility effects on happiness: Even more difficult than usually acknowledged. *European Journal of Population, 30*, 263–290.
- Margolis, R., & Myrskylä, M. (2011). A global perspective on happiness and fertility. *Population and Development Review, 37*, 29–56.
- Margolis, R., & Myrskylä, M. (2015). Parental well-being surrounding first birth as a determinant of further parity progression. *Demography, 52*, 1147–1166.

- McLanahan, S., & Adams, J. (1987). Parenthood and psychological well-being. *Annual Review of Sociology*, 13, 237–257.
- Myrskylä, M., & Margolis, R. (2014). Happiness: Before and after the kids. *Demography*, 51, 1843–1866.
- Nelson, S. K., Kushlev, K., English, T., Dunn, E. W., & Lyubomirsky, S. (2012). In defense of parenthood: Children are associated with more joy than misery. *Psychological Science*, 24, 3–10.
- Normand, S.-L. T., Landrum, M. B., Guadagnoli, E., Ayanian, J. Z., Ryan, T. J., Cleary, P. D., & McNeil, B. J. (2001). Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology*, 54, 387–398.
- Pollmann-Shult, M. (2014). Parenthood and life satisfaction: Why don't children make people happy? *Journal of Marriage and Family*, 76, 319–336.
- Pouwels, B. J. (2011). *Work, family, and happiness: Essays on interdependencies within families, life events, and time allocation decisions* (Unpublished doctoral dissertation). Utrecht University, Utrecht, The Netherlands.
- Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society: Series A*, 147, 656–666.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 3, 33–38.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Education Psychology*, 66, 688–701.
- Schoen, R., Kim, Y., Nathanson, C., Fields, J., & Astone, N. M. (1997). Why do Americans want children? *Population and Development Review*, 23, 333–358.
- Sweeting, H., Bhaskar, A., Benzeval, M., Popham, F., & Hunt, K. (2014). Changing gender roles and attitudes and their implications for well-being around the new millennium. *Social Psychiatry and Psychiatric Epidemiology*, 49, 791–809.
- Veenhoven, R. (1996). Developments in satisfaction research. *Social Indicators Research*, 20, 333–354.
- Zimmermann, A. C., & Easterlin, R. A. (2006). Happily ever after? Cohabitation, marriage, divorce and happiness in Germany. *Population and Development Review*, 32, 511–528.