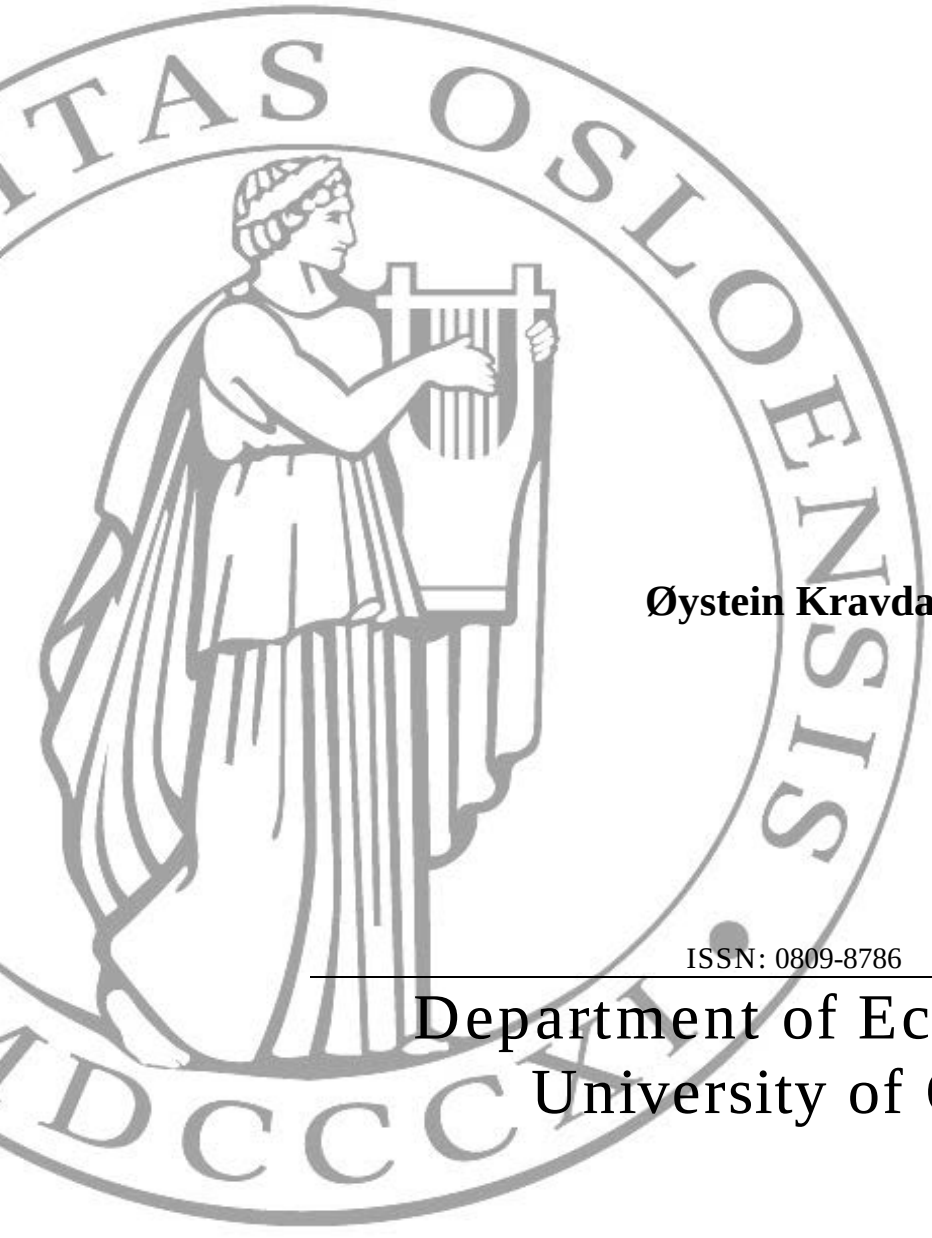
MEMORANDUM

No 10/2013

Reflections on the Search for Fertility Effects on Happiness



Øystein Kravdal

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/econ
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 09/13	Erik Biørn and Hild-Marte Bjørnsen *What Motivates Farm Couples to Seek Off-farm Labour? A Logit Analysis of Job Transitions*
No 08/13	Erik Biørn *Identifying Age-Cohort-Time Effects, Their Curvature and Interactions from Polynomials: Examples Related to Sickness Absence*
No 07/13	Alessandro Corsi and Steinar Strøm *The Price Premium for Organic Wines: Estimating a Hedonic Farm-gate Price Equations*
No 06/13	Ingvild Almås and Åshild Auglænd Johnsen *The Cost of Living in China: Implications for Inequality and Poverty*
No 05/13	André Kallåk Anundsen *Econometric Regime Shifts and the US Subprime Bubble*
No 04/13	André Kallåk Anundsen and Christian Heebøll *Supply Restrictions, Subprime Lending and Regional US Housing Prices*
No 03/13	Michael Hoel *Supply Side Climate Policy and the Green Paradox*
No 02/13	Michael Hoel and Aart de Zeeuw *Technology Agreements with Heteregeneous Countries*
No 01/13	Steinar Holden, Gisle James Natvik and Adrien Vigier *Do Medical Doctors Respond to Economic Incentives?*
No 32/12	Leif Andreassen, Maria Laura Di Tomasso and Steinar Strøm *Do Medical Doctors Respond to Economic Incentives?*

Previous issues of the memo-series are available in a PDF® format at:
http://www.sv.uio.no/econ/english/research/memorandum/

Reflections on the Search for Fertility Effects on Happiness

Øystein Kravdal

Department of Economics, University of Oslo

Memo 10/2013-v1

Contact details:

Address: P.O. Box 1095 Blindern, 0317 Oslo, Norway

Telephone (47)48020904; Fax: (47) 22855035

Email: okravdal@econ.uio.no

Reflections on the Search for Fertility Effects on Happiness

Abstract

There have been many studies of how the number of children in a family affects the parents' or the children's lives. One strand of this research focuses on the implications of fertility for the parents' level of self-reported well-being or happiness. It is argued in this paper that an overall "happiness effect" is not very informative because of the presumably large variation in individuals' perceived gains from having children. Furthermore, it is explained that such an effect would be difficult to estimate. Most importantly, the highly varying ideas about how a child will affect life quality are important for the decision about whether to *have* a child. Many of those who have few or no children have chosen this because they think their life will be best this way, and their happiness therefore tells us little about how happy their more fertile counterparts - who to a large extent have other preferences – would have been if *they* had few or no children. This estimation problem that arises because expectations about the effects of a certain behaviour (here childbearing) are heterogenous, and also affect that very behaviour, is acknowledged in the economics literature, but there is little consciousness about it in the fertility-happiness research. In addition, there is a more "standard" selection problem: factors with implications for childbearing desires, or for the chance of fulfilling these, may also affect or be linked to happiness for other reasons. Unfortunately, even the most advanced statistical approaches that have been used in this research area fail to handle all these problems, so reported results should be interpreted very cautiously.

Key words: fertility; happiness; effect heterogeneity; method; selection; subjective well-being

JEL code: J13

1. Background

There has been a strong interest in aggregate-level consequences of high fertility and high population growth (Headey and Hodge 2009; Panayoutou 1994; O'Neill 2010), and much research has also dealt with the possible societal implications of low fertility (Blake and Mayhew 2006; Dormont et al. 2006; Rand 2004). Furthermore, many studies have taken an individual- or household-level approach and addressed the implications of low or high fertility for the children in the family or the parents. For example, much attention has been given to the potentially adverse effects of large sibship size on children's education (Black et al. 2005; Eloundou-Enyegue and Williams 2006; Li et al. 2008), how the number of young children affects the parents' work activity (Boushey 2008; Dommermuth and Kitterød 2009), the importance of children as providers of help to elderly parents (Brandt et al. 2009; Wenger et al. 2007), and the emotional rewards from parenthood (Eiback and Mock 2011; Nelson et al. 2012; White and Dolan 2009). In recent years, there has also been a strong interest in how the various negative and positive effects of children on parents' lives sum up to an effect on various measures of overall subjective well-being, including what one may refer to as "life satisfaction" or "happiness". Many authors have pointed out that "life satisfaction" and "happiness" is not the same, and therefore also may be differentially affected by fertility (e.g. Haller and Hadler 2006), but it is not important in this paper to make such a distinction. All arguments simply refer to "happiness".

Some studies have shown a positive relationship between parenthood or number of children on the one hand and happiness on the other, but there are also several that have pointed in the opposite direction, and many authors have emphasized that the relationship varies with factors such as age, sex, marital status, socioeconomic status, and welfare support systems (see Aassve et al. (2012), Fritjers et al. (2004), Keizer et al. (2010), Nelson et al. (2012), Soons et al. (2009), Vannasche et al. (2012), Zimmerman and Easterlin (2006) and studies cited in a recent review by Hansen (2011)). The implications of childbearing may also depend on how early in the parents' lives the children were born, whether they are biological children or step children, the sex of the children, and whether they have had a difficult life (Umberson et al. 2010).

As in other social science areas, the underlying aim of most of this fertility-happiness research is probably to come as close as possible to identifying a causal effect, though authors may not always be so explicit about that. A causal effect may be thought of as the difference in the outcome (happiness) resulting from experiencing versus not experiencing the event (having or not having had an additional child), everything else given. It is widely acknowledged that one cannot find such an effect (or its average across several individuals, to the extent that there is variation) just by comparing, for example, two-child couples with one-child couples, because some of the factors that have led a couple to have only one child – for example a difficult economic situation or poor health - may also have implications for happiness. Researchers typically try to take this into account by controlling for such factors in various kinds of more or less advanced regression analysis. However, there is also another reason why a simple comparison with the one-child couples would be inadequate: these couples may have preferred to have only one child, and not necessarily because of a difficult

life situation, but perhaps because they have other preferences than those with two children and think their life will be best this way. Their level of happiness may well be much higher than what it would be among the two-child couples - who to a large extent have also *wanted* two children - if *they* had only one child. In fact, they may be just as happy as their two-child counterparts. Stating this in more general terms, a problem arises because one tries to estimate an effect of a factor (childbearing) that is influenced by people's expectations about that very effect – expectations that may vary greatly. This problem has received some attention in the economics literature, and ways to deal with it have been suggested (Blundell et al. 2005), but it seems not to be properly acknowledged in the fertility-happiness research. Margolis and Myrskylä (2011) very briefly hinted to it, and some authors have at least pointed out that fertility decisions are based on expected gains from childbearing (e.g. Kohler et al. 2005), but on the whole there is very little consciousness about this source of bias. In principle, there would be similar concerns when estimating effects of children on more specific outcomes, but there are special reasons to worry when the focus is on an overall measure of well-being, because this is so central when making important decisions such as whether to have a child (Haybron 2003).

The objective of this paper is to explain in a non-technical way - using simple examples and no equations - the challenges one meets when trying to assess the importance of childbearing for happiness. The first step is to justify and elaborate on the fundamental idea that many (but not all) consequences of childbearing can be foreseen and are judged differently, and are likely to be taken into account in the fertility decision-making. Then, it is discussed how this situation – which seems not to have caused any worries (despite economists' concerns about similar situations) - makes it particularly difficult to learn about the fertility effects on happiness. This is done in four steps. First, a quantitative hypothetical example is presented. It is simple, but should still represent parts of reality well enough. Second, it is discussed – with reference to this example - how meaningful it is to present an overall "effect of fertility on happiness" when there is so much variation, and it is explained how difficult it is to estimate such an effect. Third, a more "standard" type of selection mechanism – well recognized in the fertility-happiness literature – is illustrated for comparison and to build up a more complete and realistic picture. This is done through two additional examples. Fourth, it is shown more explicitly how hard it is to estimate a fertility-happiness effect – in the presence of heterogenous perceptions about childbearing consequences combined with the "standard" selection problem - even with the most advanced techniques used so far in this literature. The paper ends with a summary of the methodological challenges and a few brief reflections on the way forward.

2. Various types of consequences of childbearing and their relevance for fertility desires

Childbearing has several consequences. Most of them are probably known to people, in the sense that they are conscious about the possibility of such consequences. Their expectations about these consequences are likely to be taken into account in their fertility decision-making. Other consequences may be largely unknown. Furthermore, some consequences are rather

general, while others vary markedly – in strength and even in sign. To start with an example of the latter, most people surely realize that, if they have children, there will be periods when they will be intensely involved in care. This is probably seen as positive by many, and indeed an important reason for having children, while others may see this activity as largely a burden and prefer to do as little of it is as possible (Poortman and van der Lippe 2009). A related issue is that some may expect a strengthening of their relationship as a result of sharing responsibility for children, while others may be concerned about a possibly reduced relationship quality (Twenge et al. 2003).

Other consequences may be seen as generally positive. For example, most parents – even those who do not strongly enjoy playing with children or interacting with them in other ways – probably find parenthood rewarding in the sense that they derive emotional pleasure from seeing a child growing up, feeling needed, and knowing that they have given life to another human being who is likely survive them (Eiback and Mock 2011; Nelson et al. 2012; White and Dolan 2009). However, there are probably different views among potential parents about exactly how much this positive aspects of childbearing and -rearing would add to their happiness. Furthermore, many may have an idea that children may be helpful in old age (Brandt et al. 2009; Wenger et al. 2007), though the inclination to take such uncertain and long-term implications into account will surely also vary much, and in developing countries (poorly represented in the fertility-happiness literature so far) also younger parents benefit from practical contributions from children, who often do much household work and are strongly involved in income-generating activities. Finally, children may have high value from a religious perspective, in rich and poor countries alike (McQuillan 2004).

On the more negative side, it is probably widely expected among parents that they may have periods with little sleep while a child is young (Dørheim et al. 2009), and that there may be poorer work opportunities (Boushey 2008, Dommermuth and Kitterød 2009), more economic worries (Aassve et al. 2007), worries about the children's well-being, and less time for own leisure activities (Bittman and Wajkman 2000). In total, these burdens may lead to mental distress (Evenson and Simon 2005). However, there may be different ideas about the chance of these outcomes – depending, for example, on whether help from grandparents can be expected - and there may be different evaluations of how sleep deprivation, lower income, or a more home-based life affects life quality.

Children may exert control of health behaviour (Joutseneemi et al. 2007; Kendig et al. 2007), and parenthood may increase the level of social integration into the neighbourhood or more generally (Bühler 2008; Knoester and Eggebeen 2006; Nomaguchi and Milkie 2004). Such factors, along with some of those mentioned above, probably contribute to the inverse relationship between number of children (within reasonable limits) and mortality that has been observed for both sexes (Grundy and Kravdal 2010). The possibility of such social effects with long-term health implications is perhaps not widely considered, but it is not impossible that it enters into the decision for some people.

Reality may, of course, turn out to be better or worse than expected. In particular, the consequences that people do take into account may be stronger or weaker than foreseen. For

example, some may be more sleep deprived at the baby stage than they imagined they could be; some may draw fewer practical advantages from children in the long run than expected because they are less in need of such help, or because the children are not willing to provide support; and some may enjoy the interaction with children more than they thought originally. In the extreme, consequences thought to be positive may turn out to be negative, and vice versa. We may consider this difference between actual and expected consequences as additional unknown contributions to the actual happiness. Most of these probably vary in size and sign between individuals, but it is possible that some of them are generally positive (or negative), which would mean that people, on the whole, tend to underestimate (or overestimate) the contribution of that factor to happiness.

Furthermore, there may be consequences beyond what people usually take into account – because they are largely unknown to them. For example, there is a literature suggesting that women are influenced physiologically by pregnancies (Britt et al. 2007; Fletcher et al. 2002; Rieck and Fiander 2006; Russo and Russo 2007), with positive or negative implications for later health and thus happiness (and ultimately mortality). These findings are probably not well known to the public. There may, of course, also be other such health effects that *no one* currently is aware of, and that may be revealed in later research. This particular type of unknown effect varies perhaps quite little between people.[1]

To conclude, several consequences are likely to be taken into account, and there are potentially very different views about them, leading to large differences in fertility desires. Some think that having another child on the whole is good for them, in spite of some expected adverse implications, because the benefits dominate. Others reach the opposite conclusion, even if they perhaps are faced with very similar material conditions, because they have other lifestyle preferences. It is, of course, not easy for an individual to form an overall conclusion about the value of childbearing, because some consequences are felt soon and others only in the longer term. If childbearing, for example, is expected to – on the whole – reduce happiness during young adult years and improve it in old age (as indicated by a number of studies, including Margolis and Myrskylä 2011), it is not obvious what the overall judgment would be. In the discussion below, this timing aspect is ignored. The level of happiness is considered as constant except for an immediate change after childbirth, if any. Estimation of a "happiness effect" is problematic enough even with that simplification.

3. Example 1: simple variations in fertility, happiness and attitudes to childbearing

Let us assume that there are 300 couples who already have one child and consider having another (see Table 1). This is a highly relevant perspective, given the current below-replacement fertility in Europe. Some of the consequences of having a child that were reviewed above may be more relevant for the childless than for those who have already had a child (as indicated by Kohler et al. 2005, who found more positive happiness effects of first-born than later-born children). However, the basic picture is the same, and it is easier to discuss one-child couples than childless couples or individuals, among whom childbearing is more deeply intertwined with changes in partnership status, which also are important for

happiness (Soons et al. 2009). The timing of this second birth may also be an issue the parents consider, but that is ignored for simplicity.

Let us further assume that there are 200 couples among the 300 who think that a second child would give them (in total, as a couple) a level of happiness 5 units higher than if they remain one-child parents. Therefore, they would want a second child. (The possibility that there are not only rational ideas behind fertility intentions is not considered here.) Furthermore, let us assume that the remaining 100 couples have the opposite view: they think life will be better without a second child than if they have a second child, the difference in happiness between having and not having a child being -5. Let us also assume that other contributions to expected happiness are the same for the two groups, in the sense that the couples in the largest group expect a level of 17 if they have a second child and 12 if they do not, while those in the smaller group expect a level of 12 if they have a second child and 17 if they do not (numbers shown in the left part of Table 1). These numbers are, of course, completely arbitrary. 103 and 119 could have been chosen instead of 12 and 17. It would seem reasonable that the intensity of the views about childbearing varies, for example that some couples expect a child to change happiness by only 3 in either direction. However, this is not essential for the arguments that follow and therefore ignored. Finally, let us assume that the couples are correct in their expectations, so that the actual happiness they would attain in these situations is also 12 and 17.

This symmetry in the evaluations is meant to symbolize that one outcome is not considered generally inferior to the other, for example because it is chosen by people who are generally unhappy. People simply have different ideas about what is good for them. Those who want one child and have one child are just as happy as those who want two children and have two children, and failure to achieve the goal results in the same welfare loss for everyone. This simple example should represent well enough a part of reality, and therefore be a relevant starting point for a discussion of how problematic it is to estimate effects of fertility on happiness. Later, the example is modified to take into account that some may also have chosen to have one child because of a difficult life situation.

TABLE 1: A summary of Example 1.

NUMBER OF COUPLES	EXPECTED HAPPINESS FROM HAVING 1 OR 2 CHILDREN	NUMBERS HAVING 1 OR 2 CHILDREN IF UNRELATED TO FERTILITY DESIRES	MORE REALISTIC	ACTUAL HAPPINESS
200	Expect 17 from having 2 and 12 from having 1 -> want child	2 ch: 100 1 ch: 100	180 20	2 ch: 17 1 ch: 12
100	Expect 12 from having 2 and 17 from having 1 -> don't want child	2 ch: 50 1 ch: 50	24 76	2 ch: 12 1 ch: 17
ALL COUPLES			2 ch: 210 1 ch: 90	2 ch: 16.29 1 ch: 15.89

2 ch: 150		2 ch	15.33
1 ch: 150		1 ch:	13.67

Let us for some illustrative purposes first assume – very unrealistically - that the childbearing desire has no impact on actual childbearing: the chance of having a child is 50% for everyone, regardless of the expected consequences and thus the childbearing desires. In that case, 100 of the 200 couples in the first group will have two children and a happiness level of 17, while 100 couples will have one child and a happiness level of 12. In the other group, 50 will have two children and a happiness of 12, and 50 will have one child and a happiness of 17.

4. How meaningful is the concept of an "effect of fertility on happiness"?

The following may be seen as a reasonable definition of an "effect of childbearing on happiness": the difference in happiness that a couple would experience if they had a child compared to if they did *not* have a child, averaged over all couples. (Econometricians would typically refer to this as an "average treatment effect".) Among the first group of 200 couples in the example above, it was just assumed (unrealistically) that 100 had a child and attained a happiness level of 17. They would have attained a level of 12 had they *not* had a child. The 100 others among the 200, who had only one child and a happiness of 12, would have had a happiness of 17 had they *had* a child. In either case, the difference in happiness associated with having versus not having a child is 5. Among the second main group, consisting of 100 couples, the corresponding difference is -5. In this case, where the difference is 5 for 200 couples and -5 for 100 couples, the effect would be 1.67. This positive figure reflects that the second group is smaller, i.e. that there are fewer who think one child is the best solution than who prefer two children. It is referred to below as the "true (overall) effect", "correct effect" or something similar.

Most of the analysis of the fertility-happiness effect has been based on a simple cross-sectional design. Applied to the situation described here, the essence of such an analysis is to compare the happiness level of two-child couples with that of one-child couples. Typically, this is done in a regression approach that allows control for factors affecting both fertility and happiness, but that kind of complexity is left for later sections. The simple perspective taken now is sufficient to illustrate the problem.

The average happiness of the couples with 2 children is 15.33, and that among the one-child couples is 13.67. This gives a difference of 1.67, which is also the true overall effect as defined above. If the chance of having a child had been higher than 50% or lower than 50%, but the same in both main groups of couples, the happiness levels among two- and one-child couples would have been the same, and the difference again 1.67. More generally, if fertility were unrelated to childbearing desires, the difference in happiness between those with two children and those with one child would give us the correct effect as it is defined above.

However, it would be highly misleading to present one such number as the "effect of childbearing" when there is so much variation (childbearing being a good outcome for some and bad for others). In principle, it could in the worst case lead to unfavourable behaviour changes, because individuals informed about such an "effect" might revise their expectations about the happiness resulting from childbearing upwards, and if they were at the margin with respect to fertility desires, they might end up wanting and having a child that it would not really be in their own interest to have. Stated differently, they knew what they were doing, and information about an "effect" has not had any benefical impact. Conversely, if those who had negative attitudes to childbearing constituted a majority, the average "effect" as calculated above would be negative, and some of those barely wanting a child might be pushed to the other side against their own interests.

Some might also consider a positive happiness effect as justifying pronatalist policies: by subsidizing childbearing, so that more people have a child, the well-being of the population would increase (unless the resources allocated to this might have led to larger welfare gains elsewhere). However, this is not a good argument. Making childbearing less expensive could correspond to adding, say, 6 to the happiness for those with a child. The two groups, consisting of 200 and 100 couples, would then not rank the outcomes as 17/12 (a short-form for 17 if the couple has two children and 12 if they have one) and 12/17, but 23/12 and 18/17. The first group of 200 couples would still want a child, which would give them an even higher level of happiness, and also the second group of 100 couples would want a child, because this would now be a marginally better outcome. They would have a level of happiness 1 higher than without the policy. Obviously, one could introduce the same policy even if the first main group consisted of 100 couples and the second consisted of 200 (i.e. reversed relative size), so that the overall effect had been -1.67 instead of 1.67. The only difference would be that the total happiness gain would be smaller (100 would gain 6 from the policy and 200 would gain 1, rather than the opposite). To conclude, it is very unclear what the 1.67 estimate tells us that has any relevance for policy.[2]

To summarize, there is not one "effect", but several – actually a continuum of – effects. Ideally, some measure of this variation should be presented along with the average.

5. The reasons why estimation is problematic

As already pointed out, the assumption made above is far-fetched. In real life, those who want a child are, of course, also more likely than others to actually *have* a child. Let us assume now that couples who want to have a second child (i.e. in the first main group) have 90% probability of having that child and 10% probability of remaining one-child couples. Let us also assume that the latter outcome is a result of fecundity problems and not a downward revision of fertility desires because of, for example divorce or unexpected economic hardship (i.e. a change from a situation where everyone ranks the outcomes as 12/17 to a situation where a 10% subgroup rank them as, say, 10/8). The arguments are complex enough without such heterogeneity in attitudes to childbearing developing over time within the group. Furthermore, let us assume that 70% of those who do *not* want a child (i.e. in the second main

group) succeed in avoiding further childbearing, while 30% have an unintended birth. (This choice of 90% or 30% birth probability is not crucial for the conclusions that are drawn.) These numbers are also included in Table 1.

The average level of happiness among those with two children is, of course, higher in this situation than in the "random-childbearing" situation described in the preceding section (16.29 vs 15.33), since those who have a second child to a larger extent have wanted this because they consider it the best outcome. Also the level of happiness among those who remain one-child couples is higher than in the random-childbearing situation (15.89 vs 13.67), for the same reason: this outcome largely occurs among those who want it. The difference between one- and two-child couples is not the same, though. In this particular case, it is smaller than above (0.40 vs 1.67), i.e. we underestimate the overall effect of fertility on happiness.

While it, as explained earlier, may be rather meaningless to present an overall effect such as 1.67, it is even more meaningless to present 0.40 as the effect. This number has no straightforward interpretation. It reflects in a complex way a combination of a variety of (perceived) childbearing implications and the degree to which people actually have the number of children they want.

It should be noted that it is the combination of variation in the evaluations of the childbearing consequences and the fertility response to these evaluations that creates the problem. The implications of the latter have already been illustrated: estimation would be easy if actual fertility were not linked to the ideas about childbearing consequences. Similarly, there would be no estimation problem if all childbearing consequences were unknown, and therefore would not be relevant to take into account, regardless of whether they vary or are general. The fact that there is no estimation problem in the hypothetical situation where consequences are known and general can be illustrated by assuming that there are not two groups of couples with different attitudes, but that the first group is the only one. Then, a happiness of 17 would be observed for those with two children (who are 180 couples given the assumption about a link between fertility desires and actual fertility) and a happiness of 12 for those (20 couples) with one child, the difference and the true effect being 5.

As described earlier, the happiness associated with having a second child is, in reality, a result of a combination of childbearing consequences that vary and are known and taken into account (thus creating problems) and consequences that are largely unknown or rather general (and that would be unproblematic if they ruled the ground alone). It will now be explained that this complexity does not change the situation in any important way; estimation remains, of course, difficult.

Let us first consider the case where there is an additional effect that is both largely unknown and quite general. For example, there may as already mentioned be physiological effects of childbearing with long-term (though probably rather weak) implications for happiness (through health) that few people are aware of and that may vary little. This would correspond to adding the same number to actual happiness for everyone who has a child. If we assume that this addition is positive, more precisely +1, a comparison of happiness levels

between those with one and those with two children yields a value of 1.40 rather than 0.40. In other words, the estimate reflects the general unknown implication plus the rather uninformative measure that in a complex way summarizes subjective evaluations of implications of childbearing and the prevalence of unfulfilled fertility desires. Note that, when seeing an estimate such as 1.40, one would not be able to tell what the general effect is and what the other contribution is.

Most childbearing consequences that come in addition to those foreseen probably do vary between couples, however. Referring again to Example 1, some couples may experience a happiness bonus of 0.5, some a bonus of 1, and some a bonus of 1.5. This would mean that having a second child is generally more positive than people expect. Alternatively, excessive optimism and excessive pessimism balance each other out, so that equally many experience an additional happiness of, say, -1 and 1. Yet another possibility is that there are some couples who are clearly too optimistic and others who are more marginally over-pessimistic (e.g. additions being -3, -4, 1, and 2), or vice versa. In these situations, one would estimate an effect (i.e. a difference between the happiness of one- and two-child couples) equal to 0.40 plus the average over these unknown additional consequences.

The third possibility, in principle, is that there are additional consequences of childbearing that are known and (almost) general. That would correspond to adding a certain amount, for example 2, both to the expected happiness associated with having a child and to the actual happiness for those who do have a child. In a more complex example than described here, with a larger number of main groups having different perceptions about the consequences of childbearing, such an addition could move some groups into wanting rather than not wanting a child, but this would not fundamentally change the situation: there would be differences between the main groups in the level of happiness associated with having a child (though the average happiness gain from childbearing would be higher), and the average would tell us little and would also be mis-estimated.

To summarize, there are various types of childbearing consequences. Some are taken into account in people's decision-making, and the perceptions about them vary, while others are more general or unknown. It is the former that are responsible for the estimation problems, and that therefore are in focus also in the remaining discussion.[3]

6. Example 2: some additional realistic features incorporated

A real population does not consist only of two groups at the same overall happiness level but with opposite evaluations, such as in the preceding example. In particular, there are surely some who are generally less happy, and whose evaluation of the childbearing consequences may differ from that of others in the population. Let us now build this into the example (now called Example 2) by assuming that there are some couples who are poor and therefore also more likely than others not to have a very happy life. Obviously, there are differences in the views about childbearing also within this group of poor couples, just as among the others. Some may expect a happiness of 11 if they have one child and 6 if they have a second child,

while another subgroup of the poor would reach the opposite conclusion. In that sense, there is little new in the example. The novel feature is that whether a couple is poor or not determines the *proportions* having positive and negative attitudes to childbearing. Let us assume that those who are poor also relatively often think that a child would make the situation even worse for them.[4] More specifically, let us assume that half of the 100 couples who want to have only one child are poor, while only 10% of the 200 couples who are positive to childbearing are poor (see Table 2). Let us also assume that, regardless of poverty, the chance of having a child is 90% among those who want a child and 30% among those who do not.

TABLE 2: A summary of Example 2.

NUMBER OF COUPLES	EXPECTED HAPPINESS FROM HAVING 1 OR 2 CHILDREN	NUMBERS HAVING 1 OR 2 CHILDREN	ACTUAL HAPPINESS
180 NON-POOR			
	Expect 17 from having		
	2 and 12 from having	2 ch: 162	2 ch: 17
	1 -> want child	1 ch: 18	1 ch: 12
20 POOR			
	Expect 11 from having		
	2 and 6 from having	2 ch: 18	2 ch: 11
	1 -> want child	1 ch: 2	1 ch: 6
50 NON-POOR			
	Expect 12 from having		
	2 and 17 from having	2 ch: 15	2 ch: 12
	1 -> don't want child	1 ch: 35	1 ch: 17
50 POOR			
	Expect 6 from having		
	2 and 11 from having	2 ch: 15	2 ch: 6
	1 -> don't want child	1 ch: 35	1 ch: 11
ALL COUPLES		2 ch: 210	2 ch: 15.34
		1 ch: 90	1 ch: 13.42

This may give associations to the "standard" selection problem, in the sense that there is a factor (poverty) affecting both fertility desires and happiness, which one may therefore see a need to control for. However, a reality such as described here actually gives rise to a combination of a standard selection problem and the problem due to differences in evaluations of childbearing consequences that is dealt with in preceding sections: the example describes that people make widely different judgments, but the degree to which they land in one camp rather than the other depends on certain factors that also are important for happiness. As will be illustrated later, this means that a control for the confounding factor in a standard

regression analysis will not have the effect one might expect. An example (Example 3) of a selection problem that is not mixed with the differences-in-evaluation problem will be given in the next section.

With a reality as described in Example 2, the true effect of childbearing on happiness (as defined above) would still be 1.67, but a researcher who estimates the effect as the difference between the happiness of two-child couples and the happiness of one-child couples will get an estimate of 1.92. The reason why the estimate now is much more positive than when reality is given by Example 1 (0.40) is, of course, the very low happiness of those with one child, who to a large extent come from the subgroup who are poor, where the level of happiness is generally low.

7. Example 3: A simpler selection problem

It was assumed in Example 2 that the chance of actually having a child, given the childbearing desires, was independent of poverty. That is not necessarily reasonable. Poverty, which may be linked with low education and poor health, could be associated with infecundity or inadequate use of contraception. Rather than adding this feature to the already quite complex Example 2, let us construct a simpler example where there is no heterogeneity in childbearing attitudes. Then, we will also see an example of a "standard" selection problem not intertwined with the problem that arises from this heterogeneity and that is most central in the paper.

More specifically, assume that 200 couples fully agree that having a second child reduces their happiness by 5 (see Table 3). However, while 150 are non-poor and expect a happiness level of 17 if they avoid having another child and 12 if they have another child, there are also 50 poor couples among whom the corresponding happiness levels are 11 and 6. Let us further assume that these poor couples also are less able to avoid having the unwanted second child: 50% have a second child, as opposed to 10% among the non-poor. Then, the difference in happiness between two- and one-child couples is -7.81, which is clearly different from the true effect of -5. As shown below, this particular type of bias (if alone) is not difficult to get rid of.

TABLE 3: A summary of Example 3.

NUMBER OF COUPLES	EXPECTED HAPPINESS FROM HAVING 1 OR 2 CHILDREN	NUMBERS HAVING 1 OR 2 CHILDREN		ACTUAL HAPPINESS	
150 NON-POOR					
	Expect 12 from having				
	2 and 17 from having	2 ch:	15	2 ch:	12
	1 -> don't want child	1 ch:	135	1 ch:	17
50 POOR					
	Expect 6 from having				
	2 and 11 from having	2 ch:	25	2 ch:	6
	1 -> don't want child	1 ch:	25	1 ch:	11

ALL COUPLES	2 ch:	40		2 ch:	8.25
	1 ch:	160		1 ch:	16.06

8. Procedures that have been used in studies of the fertility-happiness effects

8.1 Cross-sectional multivariable regression

As mentioned above, most studies have essentially estimated the fertility-happiness effect by comparing the happiness level of persons who have n_1 number of children with the happiness level of persons who have n_2 number of children. The models usually control for some other factors, and many authors have additionally stratified their analysis by, for example, age or socioeconomic status or alternatively estimated the corresponding interaction effects (see e.g. Margolis and Myrskylä 2011). The main alternative to such a cross-sectional design is to do a within-individual (or within-couple) longitudinal analysis. That is dealt with below.

One may think of the cross-sectional multivariable regression analysis as producing effects of childbearing for each level of each covariate that is controlled for and then averaging these up to one effect. There will, of course, be less variation in the evaluation of childbearing consequences within a specific covariate level than in an entire national population, but there will surely be some, so the concerns that are mentioned are still valid.[5]

Let us see what happens if reality is as described in Example 2 and a linear regression model is estimated.[6] If having a second child is the only independent variable in the model, its effect is, of course, the same as the difference between two- and one-child couples mentioned earlier (1.92). If also poverty is included, however, the estimate is 0.29. To conclude, the true effect is 1.67, but because of the problem discussed first in the paper, the estimate will be 0.40. If there is also selection due to poverty, an additional bias will be introduced. The estimate changes from 0.40 to 1.92 (which happens to be closer to the true 1.67, but if another example had been chosen, the additional bias could have gone in the opposite direction). Controlling for poverty in a regression model may be considered as taking away the additional bias to some extent, as the estimate changes to 0.29, but we are not brought back to 0.40 – which would, of course, not be important either, since this estimate is of little value.

As a contrast, the estimation goes well if reality is as in Example 3: an effect of -7.81 appears, of course, in a regression model if having a second child is the only variable. If poverty is added, however, the correct effect of -5 is estimated.

8.2 Twin fixed-effects analysis

There may not be information about poverty in the data available to the researcher. However, to the extent that poverty reflects family background factors and genetic traits, it makes sense to compare among monozygotic twins. We might consider this as a special case of the cross-sectional multivariable regression approach addressed above. The idea (admittedly not very convincing) is then that twins are at the same poverty level also when they are adults. Besides, they are similar - presumably even more so - with respect to a number of other characteristics

that may also affect fertility as well as happiness. Some of these characteristics may be hard to get information about and thus control for in a standard regression.

Such an analysis has been carried out by Kohler et al. (2005), who concluded that many of the results were quite different from those obtained with standard regression. Since many probably do not have a clear intuition of the approach, it may be worth presenting it in some detail. It was assumed in Example 2 that 200 couples think (rightly) that having an additional child would increase their happiness level by 5, either from 12 to 17 or (for the few who are poor) from 6 to 11. There are also 100 who expect childbearing to reduce the happiness level by 5, from 17 to 12 or (for the much larger group of poor) from 11 to 6. Assume now that these four groups of couples are representative of one-child women in a larger population, that childbearing attitudes and poverty reflect family background factors that are shared between twins, and that the twinning rate is independent of such factors. Then, at a certain level of the twinning rate, there will be 1800 pairs of non-poor twins who agree that the happiness outcomes resulting from having/not-having a second child are 17/12, 200 pairs of poor twins who rank the outcomes as 11/6, and 500 pairs of non-poor and 500 pairs of poor twins who have the opposite view of childbearing (12/17 among the non-poor and 6/11 among the poor). We have now turned from considering couples to considering women, but that does not change the essence of the arguments. Everything discussed so far could have been based on women – or men – as the unit rather than couples without undermining the fundamental logic. The expansion from 300 couples to 3000 twin pairs is done for a very trivial reason: to avoid decimals further down in the discussion.

Let us now consider the first group of 1800 twin pairs in more detail. These women want a second child, and we assume as earlier that they have 90% probability of also having a second child. Then, there will be 18 twin pairs (1%) consisting of two women with one child, 324 pairs (18%=2·p·(1-p), where p is the individual 90% probability of having a child) consisting of one woman with one child and one woman with two children, and 1458 (81%) twin pairs consisting of two women with two children. In total, 3240 (90%) of the women have two children and 360 (10%) have one child. These numbers, which are shown in Table 4, are 20 times larger than the corresponding numbers in Example 2 (because 10 times as many twin pairs as couples are considered and because there are two women within a twin pair).

TABLE 4: A summary of the basis for the twin fixed-effects estimation

NUMBER OF TWIN PAIRS	EXPECTED HAPPINESS FROM HAVING 1 OR 2 CHILDREN	NUMBER OF WOMEN WITH ONE OR TWO CHILDREN	ACTUAL HAPPINESS OF THOSE WITH ONE OR TWO CHILDREN
1800 NON-POOR			
	Expect 17 from having	2 ch: 3240	
	2 and 12 from having	1 ch: 360	2 ch: 17
	1 -> want child	(18 twin pairs with 1+1, 324 with 1+2 and 1458 with 2+2)	1 ch: 12
200 POOR			
	Expect 11 from having	2 ch: 360	

	2 and 6 from having	1 ch: 40	2 ch:	11
	1 -> want child	(2 twin pairs with 1+1, 36 with 1+2 and 162 with 2+2)	1 ch:	6
500	NON-POOR			
	Expect 12 from having	2 ch: 300		
	2 and 17 from having	1 ch: 700	2 ch:	12
	1 -> don't want child	(245 twin pairs with 1+1, 210 with 1+2 and 45 with 2+2)	1 ch:	17
500	POOR			
	Expect 6 from having	2 ch: 300		
	2 and 11 from having	1 ch: 700	2 ch:	6
	1 -> don't want child	(245 twin pairs with 1+1, 210 with 1+2 and 45 with 2+2)	1 ch:	11

Similarly, in the second group, there are 360 who have a second child and 40 who remain one-child mothers – again 20 times the corresponding numbers in Example 2. In the two last groups, consisting of women who do not want a child and who have only 30% chance of actually having a child, there are 300 women with two children and 700 with one child (9% of the twin pairs consist of two women with two children, 42% consist of exactly one woman with one child, and 49% consist of two women with one child).

If a model is estimated for all women in these 3000 twin pairs, and only the number of children is included, the effect is of course 1.92 as in the example above. If instead a fixed-effects model is estimated by adding a dummy for each twin pair, the estimate is -0.38. If the poor twins had been exactly like the non-poor, a standard regression based on the women in the twin pairs would have given 0.40 (just as the difference between two- and one-child couples in Example 1), while a fixed-effects model again would have given -0.38. In other words, the fixed-effects approach succeeds to some extent in dealing with the bias that comes from the selection due to poverty, in the sense that the estimate is brought down from a high positive level and is the same regardless of whether there is a poor group or not. However, the other problem is not handled well: the estimate (now -0.38 rather than 0.40) is still different from the true effect of 1.67.

It is not difficult to see how the negative effect (-0.38) arises. Only the twin pairs consisting of women with different fertility contribute in the estimation. The majority of these pairs come from the twin pairs with negative attitudes to childbearing, because although the number of such twins in the population is smaller, the chance that the women in such a twin pair end up with different fertility is relatively high given the failure rate of 30% as opposed to only 10% among the others. Rather than having one group with a happiness difference of 5 and another group that is half as large and has a happiness difference of -5, as in reality (and which gives 1.67), the latter group with a difference of -5 is larger among the twin pairs who contribute in the fixed-effects analysis.

In the unrealistic situation where the chance of having a child (p) is the same among all women, regardless of their fertility desires, the difference in happiness between the woman with two children and her twin with one child is 5 for the $2000 \cdot 2 \cdot p \cdot (p-1)$ twin pairs that are positive to childbearing (recall that $2 \cdot p \cdot (p-1)$ is the proportion of the twin pairs among whom the two women have different outcomes). It is -5 for the $1000 \cdot 2 \cdot p \cdot (p-1)$ twin pairs that are

negative to childbearing. With this 2:1 relationship between the main groups, we again end up with 1.67 as the overall effect. In other words, it is the trivial fact that the differences in expected gains from childbearing have effects on actual childbearing that creates the problem, just as observed above with a simpler analysis.

An additional complicating factor when using such an approach is that there is an aspect of reality that is not reflected in the presentation so far: if the two twins have the same childbearing desires, the difference in actual fertility is a result of one of them having had one more child or one less child than wanted. That might be associated with characteristics (woman-specific and not shared between twins) that also are important for the later level of happiness. For example, an unwanted child could be due in part to modest socioeconomic resources, which could reduce happiness. Faced with such a reality, one would get a more positive estimate. Yet another source of bias is that twins may not necessarily share the childbearing preferences, contrary to the assumption above. For example, perhaps 1000 of the twin pairs actually consist of women who disagree about the implications of childbearing - one evaluating the consequences as 17/12 and the other evaluating them as 12/17.[7]

It should also be noted that a fixed-effects analysis, not surprisingly, does work well (i.e. gives the correct estimate of -5) if a real population had been as described in Example 3 and the same assumption as above about twinning rates and sharing of characteristics between twins had been made. It should be unnecessary to spell the argument out in detail. In other words, when there is homogeneity in the perceived value of childbearing and a factor shared between twins that influences the chance of actually having a child as well as the level of happiness, a fixed-effects analysis can be applied when data do not allow this factor to be controlled for in a standard regression.

8.3 Longitudinal within-individual or within-couple analysis

As mentioned, an alternative is to do a longitudinal analysis and compare the happiness of an individual after childbearing with the happiness before, while taking into account partnership formation and other changes that may occur at the same time (see, for example, Angeles 2010; Clark 2008; Fritjers et al. 2004; Zimmermann and Easterlin 2006; Soons et al. 2009; Keizer et al. 2010). The key feature of this within-individual approach is that one controls for unobserved constant individual factors that may affect both fertility and happiness (and indeed, such constant factors are important for happiness according to the so-called set-point theory; Kahneman 1999). However, an important limitation is that the data typically cover too few years to allow estimation of long-term effects. More importantly, given the perspective of this paper, also this type of analysis fails to give a correct impression of the effect when the real world is as described in Examples 1 or 2.

Let us assume that the happiness level before the potential birth of a second child can be measured, and that it is the same as assumed for the situation “remaining one-child couple”. Let us also assume that everything else is as in Example 2. Then, 162 non-poor couples would have a child and experience an increase from a happiness level of 12 to a level of 17; 18 non-poor couples would not have a child and remain at level 12; 18 poor couples would have a child and see an increase from 6 to 11; 2 poor couples would not have a child and remain at 6;

15 non-poor couples would have a child and reduce their happiness from 17 to 12; 35 non-poor couples would not have a child and remain at 17, 15 poor couples would have a child and reduce their happiness from 11 to 6; and 35 poor couples would not have a child and remain at 11. If a regression is done on the basis of the two observations of each couple, with having a second child as the only variable, its effect would be 2.50. Adding couple dummies, we instead get 3.57. If the poor had been exactly as the non-poor, the corresponding estimates would have been 2.68 and (because the approach handles the poverty selection well) 3.57.

The large effect 3.57 appears because it is essentially only those who have a child (i.e. who have experienced a change in the independent variable), and who to a large extent have wanted this, who contribute in the analysis. If, for example, the 35 + 35 couples who do not want a child and remain one-child couples *did* had a child, they would have contributed negatively (each with -5) to the happiness effect. There are also couples who have not had the child they wanted, and who would have contributed positively had they had this child, but this group is smaller (18+2 couples). It is these contributions from all groups that add up to the 1.67 effect. As with the other approaches, an effect of 1.67 would have been estimated if actual fertility were independent of the ideas about the consequences of childbearing.

Again, Example 3 can stand as a contrast. A similar longitudinal approach built on that example would give four groups of couples – one not having a child and remaining at happiness level 17, one having a child and reducing the happiness from 17 to 12, one not having a child and remaining at 11, and one having a child and reducing the happiness from 11 to 6. The effect would be correctly estimated as -5 (while a standard regression based on two observations for each couple would give -7.50).

9. Conclusions

9.1 Summarizing the problems

Childbearing is quite obviously a matter of taste to a large extent. Some expect that an additional child will give them much happiness, and may be right about that, while others – perhaps in a rather similar life situation but with other preferences – may have the opposite attitude. When the (perceived) implications of childbearing differ widely, an "effect" – defined as the population average of the difference between the level of happiness a couple would attain if they had a child and the level they would attain if they did not have a child (in econometric jargon "average treatment effect")- is not very informative. For example, a weakly positive effect may conceal the fact that some people may derive large advantages from having a child while a smaller group may gain just as much from *not* having a child.

Another problem is that it in practice would be hard to estimate such an effect. One reason is precisely this variation in the ideas about the consequences of childbearing, coupled with the trivial fact that these ideas are taken into account when making decisions. (If the consequences of childbearing were unknown, and therefore not taken into account in the decision-making, or if they varied little, there would not be similar estimation problems.) The

core issue is that most of those who have one child may have wanted this outcome and are more happy with it than the two-child couples would have been had *they* had only one child (and perhaps just as happy as the two-child couples who have wanted the second child).[8]

Another reason why estimation is difficult is that factors of importance for happiness also affect people's views about whether childbearing is good or bad for them, and that they affect the chance of having a child (through fecundity or quality of the contraceptive use) given these views about childbearing. If only the second of these problems existed, it could be solved by controlling for the relevant confounding variables, or if there is not information about these, one could control for some of them through a twin fixed-effects analysis or a within-individual fixed-effects analysis. However, even these two more advanced approaches, which are popular tools in social science these days, fail when there is variation in perceived childbearing consequences such as described here. One might say, somewhat loosely, that the procedures can handle at least to some extent the part of the problem that comes from joint determinants of fertility desires and happiness, but not the intertwined part that arises because of variation in perceived consequences, so that the former is of little help.

To conclude, results from in earlier studies (which may come from regression analysis or simpler tools, and which may or may not have been referred to by the authors as "effects") must be interpreted very cautiously. They probably do not reflect well the true overall effect, and even if they did, they could in the worst case mislead policy-makers and the public because of the underlying heterogeneity in the true effect.

9.2 Better methods or data?

In a discussion of methods that can be used to estimate effects of education on earnings, Blundell et al. (2005) pointed out that a problem arises if the effects of education differ between individuals, are known to these individuals, and are taken into account by them when they make their decisions about schooling. As argued in this paper, researchers analysing the fertility-happiness effect are in a very similar situation, though it so far seems to have been ignored: the effects of childbearing on happiness are at least to some extent known, they are heterogenous, and it is precisely these types of effects that enter into people's childbearing decisions.

Blundell et al. (2005) argued that one can deal with such a problem in an instrumental-variable approach, though one will only be able to estimate an average effect for the individuals who are induced by the instrument to take the "treatment" (have a child in our case), i.e. a so-called "local average treatment effect". It is, of course, also often difficult to find a suitable instrument. In our case, one would need a variable affecting fertility, but not (except through that channel) happiness. The authors proceed to argue that a better strategy - though also relying on such an exclusion restriction - may be to use a so-called control function method, which assumes a common unobserved factor behind the key independent variable (fertility) and the outcome (happiness). With this approach, one can estimate an "average treatment effect" – which corresponds to what has been referred to above as the true overall effect - and not only a "local average treatment effect". This latter method has been

applied also by others who have been concerned about heterogeneity in the effect of education on income (Aakvik et al. 2010).

Experimentation with such techniques is certainly one possible way to go in future fertility-happiness research, though finding a variable that satisfies the exclusion restriction is a challenge and other assumptions may also turn out to be troubling. An alterenative, and to some extent complementary, strategy would be to collect richer data. Since the core problem is that the happiness of those who have only one child is a poor measure of how a two-child couple would have fared if *they* had only one child, one might consider simply asking the two-child couples how their life would have been if they had not had their second child. However, it is not obvious that they would be able to give an answer that comes closer to the truth than an assessment based on the observed happiness of the one-child couples. Evaluating consequences of hypothetical outcomes is indeed difficult. Another approach could be to take a prospective perspective and ask younger people how they value further childbearing compared to not having more children, but that also would be problematic. While we would learn about the perceived consequences of childbearing that do influence their actual decision-making, they may be poorly informed, in the sense that what they assume to be consequences of the various outcomes might never have happened.

Faced with these obvious difficulties, would such steps even be worth trying? That is a matter of judgment. Some would say that the chance of success probably is too small to justify the costs involved in collecting the data; others would say that one would have to take some risks to be able to produce real leaps in knowledge, and that failures can give us experiences that make later efforts more successful.

That said, if reasonably good data on perceived consequences of various (hypothetical) fertility outcomes were available - which would make it possible to calculate the average as well as the variations in the gains from having a child – there would also be other important issues to address. From a welfare perspective, it is important to know how common it is to have unwanted children, or to be unable to have the child that one would like to have, and to know how much such outcomes would matter for each person's happiness. Insight into this would be relevant to those designing and scaling programs to reduce unwanted childbearing or supporting infecundity treatment. Furthermore, if would be important to find out whether people make decisions that are in their own interests, i.e. whether the actual consequences of childbearing accord with the expected. For example, if people tend to underestimate the gains from having another child (meaning that a large proportion of those with few children would be happier if they had more children), information about that should be disseminated to the population in the hope that subsequent generations will be better equipped to make good decisions.

9.3 The broader relevance of the problems discussed

The problems discussed in this paper have, of course, relevance also for studies of more specific consequences of fertility. For example, one may be interested in how childbearing affects the parents' work activity or health, or how the children's socialization (Downey and Condron 2004) or schooling careers (Black et al. 2005; Eloundou-Enyegue and Williams

2006; Li et al. 2008) are influenced by the number of siblings. To elaborate on the latter, this kind of consequence typically varies between families: possibilities for educating children vary, and people may have different ideas about how important education is for the children's and their own well-being (the former probably also being a crucial factor in the fertility decision-making, perhaps partly by being accommodated into the parents' evaluation of their own well-being). Assuming general agreement about education being important, families with many children may to a particular extent include those who think they can afford to educate so many children, and who may well be right about that. Thus, if those with a more moderate number of children had an unwanted child, there might be somewhat more adverse effects of this on the children's education than the difference between high- and medium-fertility groups would suggest. Whereas all research on how sibsize affects education takes into account - with varying degree of sophistication - that many individual and community factors may lead to high wanted or unwanted fertility as well as low education, this additional problem that subjective evaluations of implications of childbearing for education affect fertility desires is not considered.

However, there is one important difference between such studies of more specific outcomes and analyses of the fertility-happiness effect: while the more specific consequences of childbearing may vary between individuals and may be taken into account, they are not so strongly linked to variations in childbearing desires, and thus actual childbearing, as the expected consequences for happiness. Each of these consequences, such as the impact on children's education, is only one among several that contribute to the overall happiness that probably is a key factor in the decision-making.

9.4 The challenge in a nutshell

We definitely need to think more carefully about how to analyse effects of fertility on happiness. Conditions under which the existing methods work well should be better identified, and alternative approaches – including those that would require other types of data than currently used - should be considered. What we learn from this discussion may also have implications for how we analyse other micro-level effects of fertility.

10. Acknowledgements

Insightful comments from XXX and YYY are greatly appreciated. The study is part of a larger project on consequences of high fertility funded by WWW and ZZZ.

References

Aakvik A, Salvanes K.G. and Vaage K. (2010). Measuring heterogeneity in the returns to education using an education reform. *European Economic Review* 54: 483-500.

Aassve, A., Mazzuco, S., and Mencarini, L. (2006). An empirical investigation into the effect of childbearing on economic wellbeing in Europe. *Statistical Methods and Applications* 15: 209-227.

Aassve, A., Goisis, A. and Sironi, M. (2012). Happiness and childbearing across Europe. *Social Indicators Research* 108: 65-86.

Angeles L. (2010). Children and life satisfaction. *Journal of Happiness Studies* 11: 523-538.

Bittman, M. and Wajcman, J. (2000). The rush hour: The character of leisure time and gender equity. *Social Forces* 79: 165-189.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2005). The more the merrier? The effects of family size and birth order on children's education. *The Quarterly Journal of Economics* 669-700.

Blake, D., Mayhew, L. (2006). On the sustainability of the UK state pension system in the light of the population ageing and declining fertility. *The Economic Journal* 116: F286-F305.

Blundell, R., Dearden, L. and Sianesi, B. (2005). Evaluating the effect of education on earnings: models, methods, and results from the National Child Development Survey. *Journal of the Royal Statistical Society Series A* 168: 473-512.

Boushey, H. (2008). "Opting out?" The effect of children on women's employment in the United States. *Feminist Economics* 14: 1-36.

Brandt, M., Haberkern, K. and Szydlik, M. (2009). Intergenerational help and care in Europe. *European Sociological Review* 25: 585-601.

Britt, K., Ashworth, A. and Smalley, M. (2007). Pregnancy and the risk of breast cancer. *Review of Endocrine-Related Cancer* 14: 907−933.

Bühler, C. (2008). On the structural value of children and its implication on intended fertility in Bulgaria. *Demographic Research* 18: 569−610.

Clark, A. E., Diener, E., Georgellis, Y. and Lucas, R.E. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal* 118: F222–F243.

Dormont, B., Grignon, M. and Huber, H. (2006). Health expenditure growth: reassessing the threat of ageing. *Health Economics* 15: 947-963.

Dommermuth, L. and Kitterød, H.R. (2009). Fathers' employment in a father-firendly wewlfare state: does fatherhood affect men's working hours? *Community, Work & Family* 12: 417-436

Downey, D.B. and Condron, D.J. (2004). Playing well with others in kindergarten: The benefit of siblings at home. *Journal of Marriage and Family* 66: 333-350

Dørheim, S.K., Bondevik, G.T., Eberhard-Gran, M. and Bjorvatn, B. (2009). Sleep and depression in postpartum women: A population-based study. *Sleep* 32: 847-855.

Eibach, R.P. and Mock S.E. (2011). Idealizing parenthood to rationalize parental investments. *Psychological Science* 22: 203-208.

Eloundou-Enyegue, P. and Williams, L. (2006). The effects of family size on child schooling in sub-Saharan settings:A reassessment. *Demograph*y 43: 25-52.

Evenson, R. J. and Simon, R.W. (2005). Clarifying the relationship between parenthood and depression. *Journal of Health and Social Behavior* 46: 341−58.

Fletcher, B., Gulanick, M. and Lamendola, C. (2002). Risk factors for type 2 diabetes mellitus. *Journal of Cardiovascular Nursing* 16: 17−23.

Grundy, E. and Kravdal, Ø. (2010). Fertility history and cause-specific mortality: a register-based analysis of complete cohorts of Norwegian women and men. *Social Science & Medicine* 70: 1847−1857.

Haller, M. and Hadler, M. (2006). How social relations and structures can produce happiness and unhappiness: An international comparative analysis. *Social Indicators Research* 75: 169–216.

Hansen, T. (2012). Parenthood and happiness: A review of folk theories versus empirical evidence. *Social Indicators Research* 108: 29-64.

Haybron, D.M. (2003). What do we want from a theory of happiness? *Metaphilosophy* 34: 305-329.

Headey, D.K. and Hodge, A. (2009). The effect of population growth on economic growth: A meta-regression analysis of the macroeconomic literature. *Population and Development Review* 35: 221-248.

Joutsenniemi, K., Martelin T., Kestilä, L. Martikainen, P., Pirkola, S., and Koskinen, S. (2007). Living arrangements, heavy drinking and alcohol dependence. *Alcohol and Alcoholism* 42: 480-491.

Kahneman, D. (1999). Objective well-being. In D. Kahneman, E. Diener and N. Schwartz (eds.) *Well-Being: The Foundations of Hedonic Psychology*. New York: Russel Sage Foundation.

Kendig, H., Dykstra, P.A., van Gaalen, R.I. and Melkas, T. (2007). Health of aging parents and childless individuals. *Journal of Family Issues* 28: 1457−1486.

Knoester, C. and Eggebeen, D.J. (2006). The effects of the transition to parenthood and subsequent children on men's well-being and social participation. *Journal of Family Issues* 27: 1532−60.

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. and Skytthe, A. (2005). Partner + Children = Happiness? The effects of partnerships and fertility on well-being. *Population and Development Review* 31: 407-445.

Li, H., Zhang, J. and Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography* 45: 223-243.

Margolis, R. and Myrskylä, M. (2011). A global perspective on happiness and fertility," *Population and Development Review* 37: 29−56.

McLanahan, S. and Adams, J. (1989). The effects of children on adults' psychological well-being: 1957–1976. *Social Forces* 68: 124–146.

McQuillan, K. (2004). When does religion influence fertility? *Population and Development Review* 30: 25-56.

Nelson, S.K., Kushlev, K., English, T., Dunn, E.W., and Lyubomirsky, S. (2012). In defense of parenthood : Children are associated with more joy than misery. *Psyhcological Science* on line first.

Nomaguchi, K.N., and Milkie, M.A. (2004). Costs and rewards of children: The effects of becoming a parent on adults' lives. *Journal of Marriage and Family* 65: 356-374.

O'Neill et al. 2010. The impact of demographic change on carbon emissions: A global assessment. *Proceedings of the National Acadamy of Sciences*.

Panayoutou, T. (1994). Population, environment, and development nexus. In Cassen, R. (ed.). *Population and development: Old debates, new conclusions*. Transaction Publishers, New Brunswick 149-180.

Poortman, A.R. and Van Der Lippe, T. (2009). Attitudes toward housework and child care and the gendered division of labor. *Journal of Marriage and the Family* 71: 526-541.

Rand (2004). Low fertility and population ageing. Causes, consequences and policy options. Prepared for the European commission. Santa Monica CA.

Rieck, G. and Fiander, A. 2006. The effect of lifestyle factors on gynaecological Cancer. *Best Practice and Research: Clinical Obstetrics and Gynaecology* 20: 227−251.

Russo, I.H., and Russo, J. (2007). Primary prevention of breast cancer by hormone-induced differentiation. *Recent Results in Cancer Research* 174: 111-30.

Sobotka, T., Skirbekk, V. and Philipov, D. (2011). Economci recession and fertility in the developed world. *Population and Development Review* 37: 267-306.

Soons, J.P.M., Liefbroer, A.C. and Kalmijn, M. (2009). The long-term consequences of relationship formation for subjective well-being. *Journal of Marriage and Family* 71: 1254-1270.

Twenge, J.M., Campbell, W.K. and Foster C.A. (2003). Parenthood and marital satisfaction: A meta-analytic review. *Journal of Marriage and Family* 65: 574-583.

Umberson, D., Pudrovska, T. and Reczek, C. 2010. Parenthood, childlessness, and well-being: A life course perspective. *Journal of Marriage and Family* 72: 612−629.

Vanassche, S., Swicegood, G. and Matthijs, K. (2012). Marriage and children as a key to happiness? Corss-national differences in the effects of marital status and children on well-being. *Journal of Happiness Studies* online

Wenger, G.C., Dykstra, P.A., Melkas, T. and Knipscheer, K.C.M. (2007). Social embeddedness and late-life parenthood – community activity, close ties, and support networks. *Journal of Family Issues* 28: 1419-1456.

White M.P. and Dolan, P. (2009). Accounting for the richness of daily activities. *Psychological Science* 20: 1000-1008.

Zimmermann, A. and Easterlin, R. A. (2006). Happily ever after? cohabitation, marriage, divorce and happiness in Germany. *Population and Development Review* 32: 511–528.

Notes

[1] Yet another relevant issue is that parents' happiness probably depends on the quality of each child's life, which in turn may be influenced by the number of siblings he or she has, and the parents may well be conscious about the possibility of such impacts on their happiness when they consider having children. This issue is brought up again at the end of the paper.

[2] Another practical implication of a positive overall "effect" could be that efforts are made to help those who cannot have the child they want, while a negative effect might encourage attempts to avoid unwanted chilbearing. However, such steps might be warranted regardless of the overall happiness effect, as there would be welfare losses (of 5) associated with these outcomes anyway. All the 1.67 number can tell us is that the potential population-level welfare loss associated with not having a child that is wanted is larger than that associated with having an unwanted child, because a larger group wants a child.

[3] In addition to misjudging the implications of childbearing, it is of course possible that people misjudge the other contributions to happiness. This corresponds to adding or subtracting an amount to actual happiness that is unrelated to childbearing (so that the actual levels are, say, 13/18 instead of the expected 12/17). This would have no implications for the estimates of the effect of fertility on happiness.

[4] Admittedly, the literature does not provide clear evidence of an income effect on fertility desires, but low income is often linked to fear about later income decline due to for example unemployment, which is more likely to have an adverse effect (Sobotka et al. 2011).

[5] To spell out in more detail, there is variation in the (perceived) impact of childbearing, which itself can make an overall effect (within this particular covariate level) meaningless; this variation in combination with the fact that people's evaluations have a strong bearing on actual fertility creates additional problems; and there are factors behind childbearing attitudes and actual fertility above and beyond the included covariates that may affect happiness more directly.

[6] The sample is set up as in Example 2 without any additional random term, and the OLS regression module in the SAS software is used. The interest lies in the point estimates.

[7] Depending on the chance of actually having a child, some of these twin pairs would include women with the same number of children, and others would include women with different number of children. In the latter group, both women would have another number of children than wanted (i.e. both attaining a happiness of 12), or both would have the number of children they want (i.e. both attaining a happiness of 17). Having such twin pairs in the real population would produce estimates closer to 0.

[8] Further nuances could be added to the arguments. For example, some of those with one child may have originally wanted two, but experienced problems that led them to revise their

fertility desires, and may thus not be particularly happy, but still better off than if they had a second child. Also, the implications of childbearing for happiness probably vary over age, and it is not obvious whether the couples take the long- or short-term implications into account in their decision-making. However, none of these complicating factors would change the important conclusion that the situation for those with one child is not a good indicator of how the situation would alternatively have been for those with two children.