



**Zeszyty naukowe**

**Working papers**

**The family size effects on female employment.  
Evidence from the “natural experiments” related  
to human reproduction**

**Anna Baranowska-Rataj**

Working Papers  
Institute of Statistics and Demography  
Warsaw School of Economics

Nr 33, rok 2013

# **The family size effects on female employment. Evidence from the “natural experiments” related to human reproduction**

**Anna Baranowska-Rataj**

*Institute of Statistics and Demography, Warsaw School of Economics*

*Sociology Department, Umeå University*

## **Abstract**

The “natural experiments” related to human reproduction are an increasingly often applied strategy for assessment of the family size effects on female employment. The aim of this paper is to review theoretical concepts and the available empirical evidence on studies that implement this methodological approach.

Most studies confirm that the number of children does have a negative effect on female employment, net of the impact of women’s preferences regarding involvement in home-based versus paid work. Research provides consistent evidence on the way in which the effect of the number of children depends on parity and weakens over time, as the child becomes older. There is no consensus on the way that individual resources and preferences moderate the effect of family size on employment, however. Surprisingly little attention has been paid so far to the variation in the magnitude of family size effect according to living arrangements and country-specific contexts.

Keywords: family size, female labour supply, causality

JEL: J13, J22

CONTENTS

- I. INTRODUCTION ..... 4
- II. THEORETICAL CONCEPTS ..... 6
- III. METHODOLOGICAL CHALLENGES AND SOLUTIONS..... 10
- IV. MAIN FINDINGS FROM THE “NATURAL EXPERIMENTS” ..... 15
  - HETEROGENOUS EFFECTS ACROSS TIME AND SPACE* ..... 16
  - THE ROLE OF INDIVIDUAL RESOURCES AND PREFERENCES*..... 21
  - DURATION OF THE EFFECTS* ..... 22
- V. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH..... 23
- ACKNOWLEDGEMENTS ..... 24
- REFERENCES ..... 24

## **I. INTRODUCTION**

The impact of childbearing on female employment is important for a number of reasons, and the primacy of these reasons depends on the country-specific context. For the developing countries, a strong negative causal relationship between childbearing and female employment means that family planning programmes could possibly raise female labour force participation and hence foster economic growth (Schultz 2008; Cristia 2008). In the developed countries, whose societies are ageing, the evidence on the effects of childbearing indicates whether the two key policy goals: raising fertility - that helps to keep a balance between the generations, and increasing female employment - that alleviates the negative consequences of the ageing of the workforce, are contradictory (Matysiak & Białowolska-Węziak 2012). Finally, understanding the impact of childbearing on female labour market outcomes is crucial for all policymakers who aim at improving the gender equality.

This paper serves a complementary role to the currently available reviews of the literature on the relationship between fertility and female employment (Matysiak & Vignoli 2008; Del Boca & Locatelli 2002; Browning 1992). It extends the existing literature twofold. First of all, unlike most previous studies, we concentrate explicitly on the role of the tensions between motherhood and paid work that are intensified due to increased family size after the birth of any additional child. Second, we discuss the methodological challenges related to evaluation of the causal effects of family size on female employment and we present the evidence from studies that meet these challenges by applying a quasi-experimental design. The use of such research design has been so far rather uncommon in demographic research, however, has been receiving increasing attention in population studies (Moffit 2003, 2005).

The existing literature focuses on the maternity break and the return to work. On the one hand, as children become older, they require less time and attention, thus making the tensions between motherhood and occupational roles less severe. This means that the return to work becomes easier across the time that passes from the childbirth. On the other hand, the

prolonging maternity break causes human capital losses and thus lowers the employment chances. The dynamics of the return to work after a birth are a joint outcome of these two processes: decreasing time conflict between work and family duties and human capital depreciation.

The foci of this paper are not the transitions to work after childbirth, but instead the impact of the number of children on women's employment. From the conceptual point of view, this is a different research question. Increasing the family size multiplies the time that has to be invested in parenthood duties. Hence, rather than looking at the prolonged transition to work after childbirth, we look at the way that each additional child decreases mother's employment chances. The empirical assessment of the magnitude of this decrease is difficult due to the mechanisms behind family decision-making. In many modern societies, most births do not occur randomly, but rather are planned by couples. The plans regarding family size depend on a whole set of factors, including preferences regarding family size and women's family orientation (Cigno 1991; Rosenzweig & Wolpin 1980; Hotz et al. 1988). The preferred family size tends to be correlated with the preferences regarding labour market participation, with women oriented at larger family being less motivated to work. Therefore, it is quite difficult to separate the genuine impact of raising a larger number of children on female employment from the effect of the preferences regarding involvement in paid work.

These problems have been evident for social scientists for a long time, and still, the family size effects on female employment has been so far examined by means of standard regression methods (Matysiak & Vignoli 2008). These methods allow to control for a range of important factors, such as women's education and working experience obtained prior to childbearing or family income, but do not allow to take unobserved female characteristics, such as preferences, into account. Hence the results from this strand of research usually mix up the genuine family size effects on female employment and the effect of individual preferences. Such findings may be very interesting from descriptive point of view, but provide little contribution

into the academic debate on the causal consequences of childbearing and cannot be treated as guidance for policy makers.

Recently, there has been a lot of discussion on using alternative approaches that may provide insight into the causal effects of family size on female employment. Actually, this shift in emphasis from descriptive research to causal analysis can be observed in the whole discipline of population studies, not just in the field of fertility and female employment (Engelhardt et al. 2009; Moffit 2005; Schutz 2008). Among the methods that may be considered, one can distinguish natural experiments, which exploit random, sudden and unpredictable changes in environmental or personal traits (Angrist & Pischke 2010). Such quasi-experimental approach is argued to maximize internal validity. In other words, as compared to other methods, it provides most confidence that the estimates are unbiased (Moffit 2003). In this paper we provide a detailed overview of studies that draw on natural experiments in order to disentangle the causal effects of childbearing.

This paper is structured in the following way. First, we describe the theoretical concepts behind the analyses of the impact of the number of children on female employment. Next, we elaborate on the methodological challenges related to testing empirically the hypothesis on such impact. Moreover, we discuss the solutions that have been proposed to tackle endogeneity of family size with respect to female labour market attachment. Then, we provide a synthesis of the main findings from research that has implemented natural experiment approach. We summarise the results on the magnitude of the effect of having an additional child on female employment across the socio-economic groups of women and differential contexts. Finally, we present the conclusions from this literature overview and suggest the directions for future research.

## **II. THEORETICAL CONCEPTS**

The relationship between family size and female employment is very well grounded in sociological literature. This strand of research stresses that due to a number of cultural and economic factors, mothers bear the primary responsibility for childcare (Lehrer & Nerlove

1986). Both work-related duties and childcare may provide women with rewards and satisfaction, but due to time constraints, women need to decide how to best divide their time between paid work and taking care of children. This notion has been referred to as the role incompatibility hypothesis (Brewster & Rindfuss 2000).

Similar concerns have been raised in the neo-classical economic models. The economic literature on the impact of childbearing on female labour supply incorporates the key insights from the models of the demand for children, and combines them with theoretical concepts of the life-cycle human capital accumulation process (Mincer 1970; Mincer & Ofek 1982) and endogenous labour supply (Weiss 1972; Heckman 1976; Weiss & Gronau 1981). According to the theory on the demand for children, parents derive benefits from spending the time with their offspring (Lehrer, Grossbard-Shechtman, and Leasure 1996). Raising children requires financial resources, but also implies indirect costs of the foregone earnings of the person caring for the child in the home, usually the mother. Hence, a decision to have (another) child implies additional expenditures from the household budget and in the same time reduces income opportunities. Therefore, the decisions about the number of children that parents want to have and the decisions on each parents' time allocation in paid work versus childcare need to be taken jointly.

The neo-classical economic literature emphasizes the trade-off between family size and parental – usually mothers' - employment. In particular, two mechanisms leading to a decrease of the women's labour market supply after each birth can be indicated (Willis 1973; Cain & Dooley 1976; Rosenzweig & Wolpin 1980). First of all, women devote their time to taking care of children instead of doing paid work and hence do not accumulate new job-related experience. Second, the career break causes depreciation of skills that they have already gained before the birth of the child, and this additionally depresses their future employment chances.

The early work on the impact of childbearing on female employment made a number of simplifying assumptions. First of all, it tended to adopt a static viewpoint on the mothers'

involvement in childcare and investments in labour market-specific human capital, even though the decisions on these investments should be regarded as a life-cycle process (Namboodiri 1972; Heckman & Willis 1976; Moffit 1984). Second, it assumed that the goods and services consumed at home do not contribute to the child quality and that children benefit only from the time that mothers spend at home with them. Third, these models did not take into account that women may actually buy childcare services. Finally, the theoretical models linking the demand for children and female labour market career choices treated individual preferences of women as factors that do not evolve over the life course.

More recent developments in the literature have been oriented towards taking a dynamic perspective on life cycle fertility and female labour supply (Moffitt 1984; Hotz & Miller 1988; Francesconi 2002). Rather than assuming that the working experience accumulated in the pre-childbearing period increases the average lifetime wage thus raising the value of a potential mother's time, these models consider that the already accumulated human capital at each specific point in time affects the potential wage in and thus affects the cost of having another child. In other words, the decisions made at each specific age depend on the fertility and employment histories up to that point in time. Hence, these models examine the dynamic feedback effects in the relationship between fertility and labour supply with endogenously accumulated work experience.

The extensions of the basic models of demand for children and female labour supply have also considered the role of child care costs (Hotz & Miller 1988; Michalopoulos et al. 1992; Ermisch 1989) and availability (Haan & Wrohlich 2011). Interestingly, according to this research, reducing the prices of child care services does not necessarily increase female labour supply (Ermisch 1989). A lower price of child care decreases the direct costs of children, raising the demand for children and for mother's involvement in child care, which actually may lead to reduction in female employment. This mechanism is not necessarily relevant on the markets where the physical availability rather than the price of childcare services is the main obstacle to

“outsource” childcare or where the childcare subsidies are conditional on employment (Wronlich 2011; Haan & Wrohlich 2011).

Most theoretical models on the demand for children and female labour force supply focus on the causal effects of family size on women’s career chances while treating preferences of women as factors that distort any conceptual analysis. Still, it is worthwhile to discuss the direction in which these individual preferences may actually work. The crucial distinction concerns the tastes for activities in which maternal time can be invested: paid work, childcare and leisure. Some theoretical models assume for simplicity that women do not distinguish between leisure and taking care of children (Michalopoulos et al. 1992). In such setup, it could be expected that women with the highest earnings profiles, that implies a comparative advantage in market work, may display low marginal utility of children; whereas family-oriented women may have low preference for doing paid work. This implies the negative correlation of preferences towards labour market career and family career (Francesconi 2002; Lehrer & Nerlove 1986). In principle it could be also argued, however, that women treat home-based duties similarly to duties related to paid work. In that case, one should distinguish a taste for leisure from the preferences for doing any type for work (whether home based or market based). In such case, the preferences towards larger number of children could be actually positively correlated with motivation for paid work.

An important insight from the literature concerns also the stochastic nature of the individual preferences. Most theoretical and empirical work has so far assumed that the preferences are stable in time. However, it needs to be noted that the tastes for paid work, childcare and leisure may vary across various life phases. Specifically, the presence of children in the family significantly affects female preferences for these three types of activities (Joshi 1998; Matysiak 2011). Hence, after each birth, and especially after the first one, that marks transition to parenthood, individual preferences may actually change.

### **III. METHODOLOGICAL CHALLENGES AND SOLUTIONS**

Given that decisions on childbearing and employment may have common unobserved antecedents, such as individual preferences, comparisons of employment of mothers with one child and those with more children by means of standard regression models might lead to misleading conclusions. Still, the available empirical studies have usually employed models that assume exogeneity of the number of children, such as linear or logistic regression models and standard duration models (Matysiak & Vignoli 2008).

More recent studies take advantage of longitudinal data and consider the influences of unobserved factors such as individual preferences. These methods include hazard models controlling for unobserved heterogeneity, panel data models with fixed effects and models with instrumental variables (Matysiak & Vignoli 2008). Each of the above mentioned approaches has some advantages and disadvantages, and elaborating them in detail goes beyond the scope of this paper. Hazard models controlling for unobserved heterogeneity require assumptions on the distribution of the unobservables and assume that these factors are fixed in time. Fixed effect models do not require the assumptions on the distribution of the unobserved factors, but they only handle unobserved heterogeneity that does not change across the time. This is problematic since, as mentioned in the previous section, individual preferences regarding family and work may evolve over the life course.

Preferably, we would like to observe how female labour market outcomes respond to an exogenous variation in the number of children within a family (Jacobsen, Pearce & Rosenbloom 1999). If we had a variable that randomises women according to their family size, we could handle not only the time-fixed but also time-varying factors which may confound the estimation results. In life sciences an exogenous variation can be easily provided in a laboratory setting, but experiments are hardly feasible in population studies. However, one can draw on natural experiments, which exploit random, unpredictable changes in environmental or personal traits

(Angrist & Pischke 2010). These changes may be used to construct instrumental variables in simultaneous equations models.

As compared to other methods, “natural experiment” approach provides most confidence about lack of systematic bias in the estimates (Moffit 2003). The common criticism is that in the quest for internal validity, studies following this approach have become idiosyncratic (Rosenzweig and Wolpin 2000; Heckman and Urzua 2009; Deaton 2009). Indeed, “natural experiment” approach gives very limited opportunity for extrapolating the results. The more specific the treatment, the stronger the exogenous shift in the family size that we can study, and the greater internal validity of our estimation. However, this comes at the cost of external validity, which is closely related to generalising the conclusions from this estimation beyond the subpopulation that can be regarded as the “treatment” group in our “natural experiment” (Moffit 2003). In defence of this approach, Angrist and Pischke (2010) emphasize that the empirical evidence on any given causal effect is always derived from a particular time, place, and research design. As a solution to this problem, they propose to accumulate the empirical evidence, so that a more general picture begins to emerge.

The early attempts to use the exogenous variation in the number of children within a family relied on instruments such as religion or ethnic group, the mother’s number of siblings, the mother’s opinions on ideal family size, and duration of marriage (Browning 1992). However, in fact the arguments about the lack of correlation of these variables with labour market outcomes are rather difficult to defend. Research has therefore shifted towards exploiting idiosyncratic changes in policies as sources of exogenous variation in childbearing. These include reforms that affect individuals’ work incentives or tax credits (Blundell et al., 1998), or programmes related to contraception availability in the community (Arpino & Aassve 2013). However, this approach gives the opportunity to measure the family size effects only in selected countries – namely in those that actually did implement some specific reforms.

A recent and burgeoning strand in the literature exploits other type of “natural experiments”, which are associated with human reproduction and hence occur in virtually any country in the world. Specifically, these studies exploit data on miscarriages, presence of fecundity problems, multiple births as well as gender composition of offspring. Instrumental variables on miscarriages and fecundity problems allow to compare the labour market outcomes of childless women and mothers with one child, whereas instruments based on multiple births and gender composition of siblings allow estimating the effects of births at the second and third parity, respectively. In the following, we will describe the general idea of these research designs, their advantages and disadvantages.

In order to identify the effect of experiencing a birth of a child, studies have exploited the fact that some women who become pregnant, experience a miscarriage and thus do not give birth at the time they planned to become mothers. Some miscarriages might occur at random due to formation of abnormal fetal chromosomes at the time of conception, which causes fetal expulsion early in a pregnancy (Hotz et al. 1997; Hotz et al. 2005). Therefore, miscarriages can be treated as an instrument in order to obtain unbiased estimates of the causal effects of childbearing on women's labour market outcomes. Still, the validity of the instrumental variable based on data of miscarriages is a very controversial issue. First of all, the data on miscarriages usually come from the surveys that collect self-reported information, which may lead to underestimating the range of the phenomenon (and this underrepresentation proceeds in a selective way). Moreover, epidemiological studies have found that the consumption of cigarettes and alcohol during pregnancy significantly increase the incidence of miscarriages (Kline et al. 1989). In the same time, smoking cigarettes and drinking alcohol are likely to be correlated with the labour market outcomes (Bray 2005; Johansson et al., 2007; Levine et al. 1997). Since the group of women who miscarry constitute a mixture of women who experience random and non-random miscarriages, where it is typically unknown who experienced which type of miscarriage, one can argue that data on miscarriages can be regarded as a “contaminated sample”. Even

though identifying causal effects of childbearing on labour market outcomes based on such data is not possible, it is still possible to calculate bounds on this effect (Hotz et al. 1997).

The second measure that can be used to identify the labour market effects of childbearing is information on the infertility problems. The medical literature defines infertility as the failure to conceive after a year of regular intercourse without contraception (Habbema et al. 2004). The examples of applications are studies by Agüero and Marks (2008, 2011) and Cristia (2008), who have used the data on self-reported fecundity problems in order to assess the causal effects of childbearing on labour market outcomes in the developing countries. It cannot be asserted that fecundity problems are random. Infertility is known to increase as women age, but of course age can be easily controlled for. However, there is a whole range of other factors such as poor health, smoking, drinking and extreme body mass index, that are associated with infertility and may depress labour market chances. Agüero and Marks (2008, 2011) argue that the evidence on these correlates comes from couples studies where the data were not drawn from random samples of population and hence aren't reliable. Nevertheless, it remains an open question if the health-related behaviours may simultaneously determine both fecundity and women's labour market performance, confounding the estimates obtained from such "natural experiment".

Another type of "natural experiment" that shifts the family size is the occurrence of multiple births. Rosenzweig and Wolpin (1980) have been the first to propose to use the data on multiple births in order to construct a "control group" for parents with a given number of children. Multiple births are essentially an outcome of a random process and not a result of deliberate decisions or a consequence of health-related problems. They raise a higher number of births within a family, but do not affect directly the labour market performance of parents. Thus, information on twin births can be applied to construct an instrumental variable and to get unbiased estimates of the impact of the number of children on female labour market outcomes.

The approach proposed by Rosenzweig and Wolpin (1980) has been regarded as very promising approach in research on the causal effects of fertility and has actually been followed by many other studies (see e.g. Caceres-Delpiano 2006, 2012; Jacobsen et al. 1999; Karbownik & Myck 2012; Vere 2011). Still, it does have some drawbacks. First of all, it does not allow us to measure the effect of the change in the number of children from zero to one child. Following this approach gives the opportunity to measure the family size effects only at higher parities only. Second, the occurrence of multiple births correlates with some demographic variables such as age or race (Martin and Park 1999). The age at first birth can be controlled for in instrumental variable regression models, though. A more problematic issue is that some of the assisted reproduction technologies tend to lead to multiple births, and it cannot be excluded that receiving such treatment is related to some female characteristics that affect female employment. The information on infertility treatment is hardly ever available for social scientists. However, while fertility treatments do increase twinning rates, these effects are overwhelmingly concentrated among women over age 35, who are both more likely to use these treatments and more likely to have twins as a result of the treatments (Martin et al., 2003). Moreover, twins have been shown to have specific health characteristics, e.g. lower birth weight. The children born in multiple births may therefore require relatively more maternal care, as compared to children born as singletons. Finally, raising children born in multiple births may affect labour market outcomes differently than raising other children due to economies of scale. For example, since children born in multiple births often attend the same classes, parents need to spend relatively less time on helping them with homework (Rosenzweig & Zhang 2009). These arguments pinpoint the limitations of internal and external validity of instrumental variables constructed based on multiple births.

Another “natural experiment” that gives the opportunity to study the family size effects on labour market outcomes is the siblings gender composition. There is well established evidence on a preference for ‘balanced’ families with equal numbers of boys and girls in developed countries (Andersson et al. 2006; Pebley and Westoff 1982; Williamson 1983).

Angrist & Evans (1998) have been the first to exploit this kind of exogenous variation in family size, and many studies have followed this strategy (Cruses & Galiani 2007; Daouli et al 2009; Nam 2010; Karbownik & Myck 2012). Apart from the problem that this approach can only reveal the effects of shifting the family size beyond parity two, there is a debate on the potential problems with the internal validity of the instrumental variable constructed based on information of siblings gender composition. Although a child's sex is random, in some developing countries, the gender of children could have a wealth effect on female employment (Aguero & Marks 2011; Schultz 2008) and this might have a direct effect on parental labor supply. It might be also problematic to use this approach in societies where parents exhibit a strong preference toward children of a particular sex and selective abortion of female fetuses is practiced (e.g. Das Gupta, et al. 2003). Another sort of problems has been pointed out by Rosenzweig and Wolpin (2000) and refers to "hand-me-down" effects" in raising children. Same sex siblings are related to substantially lower levels of expenditures, which are more likely to arise when there are children of the same sex in the household for items such as clothing. Since these items represent a sizeable fraction of the household's expenditures, they note that the sex composition of children plausibly alters labour supply through mechanisms other than through the number of children alone. If households with children of the same sex spend less money than do households with children of differential gender sexes, this may affect the labour supply of married women.

#### **IV. MAIN FINDINGS FROM THE "NATURAL EXPERIMENTS"**

The evidence from about twenty studies existing in the literature that followed a "natural experiment" approach has been summarised in Table 1. These studies use large datasets, usually census data or large scale surveys, such as Demographic and Health Surveys. Most of these studies have been carried out in the United States or in the developing countries; while surprisingly little evidence comes from Europe. On one hand, that despite geographic diversity of the countries, whose data has been so far used in the "natural experiments", reviewed here,

most of these countries have rather unfavourable institutional arrangements for combining work and family duties. Given lack of reconciliation policies implemented in these countries, these studies cannot be really regarded as a conservative test of hypothesis on the negative impact of the number of children on female employment chances. On other hand, in most of these countries generous family benefits are missing. Hence, it remains an open question if and how the magnitude of the family size effects varies depending on the country-specific institutional context.

Many studies reviewed in Table 1 show a negative effect of childbearing on female labour supply. The magnitude of this effect varies strongly across studies.<sup>1</sup> Clearly, it depends a lot on the parity: studies that consider the effect of increasing the family size from zero to one child suggest a very small effect (cp. Agüero and Marks 2008, 2011), already a bit larger effects can be observed in studies that look at the impact of moving beyond second parity (Bronars and Grogger 1994; Nam 2010; Jacobsen, et al. 1999) and much larger effects tend to be reported in studies that consider births of third and higher orders (e.g. Daouli et al. 2009; Cruces and Galiani 2007; Angrist and Evans 1998). Very few studies report positive effects. Iacovou (2001) provides some evidence that in UK having a third child results an increase in the probability of paid work among women, but these effects are insignificant (and are commented with some reservations by the author). Rosenzweig and Wolpin (1980) show positive effects among older women. Caceres-Delpiano (2012) notes that the effects of second birth are positive for some samples and types of job. However, in general, most studies suggest that increasing the family size is either neutral or negative for female employment.

#### *HETEROGENOUS EFFECTS ACROSS TIME AND SPACE*

Opportunities for drawing broader conclusions on heterogeneity in the effects across time and space are limited in this paper due to rather restricted range of the studies carried out

---

<sup>1</sup> The reported effects vary across studies depending on the type of instrumental variable used, partly due to the mechanisms explained here. For example, instrumental variable based on twins necessarily considers children that are relatively older than children in the IV based on siblings gender (Angrist & Evans 1998).

so far. Having said that, one can still note that in US alone the causal effect of the family size becomes weaker over time. The evidence provided by Angrist & Evans 1998, Bronars and Grogger (1994), Jacobsen et al. (1999) and Vere (2011), shows a weakening impact of family size on employment. A question arises if we could ascribe this change to the institutional, structural or cultural changes. Recent studies argue that US have lagged behind other developed countries when it comes to implementing family-friendly policies (Blau and Kahn 2013), but perhaps the diminishing causal effect of family size can be attributed to improving attitudes towards gender equality or changes in the structure of jobs towards ones that are more easily combined with family duties.

Comparative studies suggest that there are differences in the family size effects between high income and low income countries, with stronger negative impact observed in the latter group (Aguero & Marks 2011). Again, it remains unclear whether such a diversity of effects can be observed due to institutional or cultural differences across countries. It seems that European countries would be a better laboratory for exploring these differences because these policy and culture-related differences are adequately described and measured (see e.g. Matysiak & Weziak-Bialowska 2012, Javornik 2012).

Table 1. The effects of the number of children on female employment.

<b>Authors, Journal</b>	<b>Instruments</b>	<b>Data</b>	<b>Family size effect on female employment (in percentage points)</b>
Agüero & Marks (2008), The American Economic Review	infertility	Demographic and Health Surveys 1994-1998	standard regression: family size decreases employment by 3.2 p.p. IV models: family size decreases employment by 0.003 p.p. (statistically insignificant effect)
Agüero & Marks (2011), Journal of Human Resources	infertility	Demographic and Health Surveys 1994-1999	standard regression: family size reduces employment by 2.4 p.p. IV models: family size effect is a statistically insignificant 0.006 p.p.
Cristia (2008), Journal of Human Resources	infertility	US National Survey of Family Growth 1973- 2002	standard regression: - IV models: having a child younger than one-year-old decreases female employment by 28 p.p.
Xia (2010), unpublished	miscarriage	US National Survey of Family Growth	standard regression: the effect of the number of children is a reduction in employment by 7.5 p.p. IV models: the effect of the number of children is a reduction in employment by 2.4 p.p.
Rozenzweig & Wolpin (1980), Journal of Political Economy	multiple birth at first birth	US National Fertility Survey 1965, National Survey of Family Growth 1973	standard regression: the effect of 2 <sup>nd</sup> child for women aged 15-24 is the decrease in employment of 0.103 p.p.; effects are insignificant in other age groups IV models: the effect of 2 <sup>nd</sup> child: for women aged 15-24, the decrease in probability of working amounts to 0.371 p.p.; the effect is 0.102 p.p. for the 25-34-year-olds; women aged 35-44 have employment chances <i>greater</i> by 0.142 p.p.
Bronars & Grogger (1994), The American Economic Review	multiple birth at first birth	US Census 1970, 1980	standard regression: ??? IV models: 2 <sup>nd+</sup> child reduces employment by 2.3 p.p. in 1970 and 1 p.p. in 1980, the effects are insignificant
Jacobsen et al. (1999), Journal of Human Resources	multiple birth at first birth	US Census 1970, 1980	standard regression: - IV models: the effect of 2 <sup>nd+</sup> child: a decrease in employment by 1.4 p.p. in 1969 and by 1.6 p.p. in 1979
Vere (2011), Oxford Economic Papers	multiple births	US Census 1980, 1990, 2000	standard regression: - IV models: 2 <sup>nd+</sup> child leads to employment decline of 11.2 p.p. in 1980, 12.6 in 1990 and 12.9 in 2000; the effect of 3 <sup>rd+</sup> child is 8.6 p.p. in 1980, 9.5 p.p. in 1990 and 7.8 in 2000
Caceres-Delpiano (2006), Journal of Human	multiple births	US Census 1980	standard regression: 10.7 reduction in employment due to 2 <sup>nd+</sup> child and 8.7 p.p. due to 3 <sup>rd+</sup> child

Resources			IV models: as a result of 2 <sup>nd+</sup> child employment decreases by 5.3 p.p.; the effect of 3 <sup>rd+</sup> child is an employment decrease of 4.5 p.p.
Cáceres-Delpiano (2012), Demography	multiple births	Demographic and Health Surveys 1994-2006	standard regression: 2 <sup>nd+</sup> child decreases employment by 1.43 p.p. , 3 <sup>rd+</sup> child by 1 p.p. (insignificant effect), 4 <sup>th+</sup> child by 0.88 p.p., 5 <sup>th+</sup> by 0.81 p.p. IV models: 2 <sup>nd+</sup> child decreases employment by 2.97 p.p. , 3 <sup>rd+</sup> child by 1.63 (insignificant effect), 4 <sup>th+</sup> child by 2.19 p.p., 5 <sup>th+</sup> by 3.81 p.p.
Carrasco (2001) Journal of Business and Economic Statistics	siblings gender composition	US Panel Study of Income Dynamics (PSID) 1986-1989	standard regression: family size reduces employment by 7.1 p.p. IV models*: family size causes an employment decrease by 12.9 p.p.
Angrist & Evans (1998), The American Economic Review	siblings gender composition	US Census 1970, 1980	standard regression: 3 <sup>rd+</sup> decreases employment by 17.6 p.p. in 1980 and 15.5 p.p. in 1990 IV models: 3 <sup>rd+</sup> child leads to a 11.3-12 p.p. decrease in probability of work in 1980 and 9.2 p.p. decrease in 1990
Iacovou (2001), unpublished	siblings gender composition	British Household Panel Study, National Child Development Study	standard regression: the effect of 3 <sup>rd+</sup> child is a reduction in the probability of work of between 12.8-14.9 p.p. IV models: having a 3 <sup>rd+</sup> child results an increase in the probability of going out to work, of between 7-13 p.p. (estimates is imprecise, effects may in fact equal to zero)
Chun & Oh (2002), Applied Economics Letters	siblings gender composition	Korean National Survey of Family Income and Expenditure	standard regression: the effect of 3 <sup>rd+</sup> child is a reduction of 5.7p.p. IV models: the effect of 3 <sup>rd+</sup> child is a reduction in employment by 27.5 p.p.
Foley & York (2005), unpublished	siblings gender composition	US census 1950, 1960, 1970, 1980, 1990, 2000	standard regression: the effect of 3 <sup>rd+</sup> child is 14.22 p.p. IV models: 3 <sup>rd+</sup> child reduces employment by 7.99 p.p.
Cruces & Galiani (2007), Labour Economics	siblings gender composition	Mexican 2000 and Argentinean 1991 censuses	standard regression: 3 <sup>rd+</sup> child reduces female employment by 9.7 p.p. in Argentina and by 9 p.p. in Mexico IV models: 3 <sup>rd+</sup> child reduces female employment by 6.5-8.1 p.p. in Argentina and by 4.4-6.3 p.p. in Mexico
Daouli et al. (2009), Economics Letters	siblings gender composition	Greek Census 1991, 2001	standard regression: 3 <sup>rd+</sup> child leads to a reduction in employment by 8.3 in 1991 and 10.5 percentage points in 2001 IV models: 3 <sup>rd+</sup> child reduces employment by 13.6 p.p. and 10 p.p.,

			respectively
Nam (2010), Labour	siblings gender composition	Korean Census 1980, 2000	standard regression: positive effect of 3 <sup>rd+</sup> child: 2.6-4.5 p.p IV models: 3 <sup>rd+</sup> child reduces employment by about 2.7-4.4 p.p.
Karbownik & Myck (2012), unpublished	multiple births and siblings gender composition	Polish Household Budgets' Survey 2003-2010	standard regression: 2 <sup>nd+</sup> child reduces employment by 8.3 pp., the negative effect of additional children in the sample of mothers with at least two children is 6.8 p.p. IV models: 2 <sup>nd+</sup> child reduces employment by 7.1 p.p.; among mothers with at least 2 children the effect of additional child is insignificant (both in case of using the twinning and siblings gender composition instrumental variable)

Note: We denote the effect of second and subsequent child as 2<sup>nd+</sup> and accordingly, the effect of third and subsequent child as 3<sup>rd+</sup>. Standard regression means regressions that do not take endogeneity of childbearing decisions into account (i.e. OLS, with exception of Rosenzweig & Wolpin (1980) and Chun and Oh (2002) who use a logit and probit model, respectively). By IV models we mean here linear regression models with instrumental variables or selection models which exploit multiple births or siblings composition to construct control function. \* Carrasco (2001) combines fixed effects modes estimated on panel data with instrumental variable approach.

*THE ROLE OF INDIVIDUAL RESOURCES AND PREFERENCES*

There are theoretically based reasons to expect a diversity of effects of family size according to women with different individual resources. For example, among more educated women, labour supply should be more sensitive to increases in family size than among less educated (Angrist & Evans 1996b). Some studies provide evidence that support such theoretical predictions. For example, Karbownik and Myck (2012) Cáceres-Delpiano (2012) find that in Poland and in the developing countries, respectively, the strongest negative effects of family size can be observed among better educated women. However, Angrist & Evans (1998) show that in US actually women with relatively low levels of schooling experience the largest effects of children on labour supply. It remains an open question, therefore, whether there exists uniform educational gradient in the magnitude of the family size effect on female employment.

From the studies quoted in Table 1 one can conclude that typically, the estimates from the “natural experiments” show much lower impact of children than the standard regression results obtained with the same data and with the same model specification. Hence, much of the negative effects observed in descriptive studies on the influence of family size on female labour market chances is simply due to selection bias. In other words, the descriptive studies capture the genuine effect of childbearing on women’s employment plus the effect of their preferences regarding labour market attachment. The direction of this bias seems to confirm the theoretical concepts derived from the sociological and economic theories: women, who have greater earnings potential, usually tend to have fewer children. In the same time, women who do not have strong advantage in terms of the labour market chances, tend to have more children.

There are some exceptions to this seemingly common pattern of work and family choices, though. Specifically, this applies to the findings obtained by Daouli et al. (2009) for Greece and Chun and Oh (2002) for Korea as well as Cáceres-Delpiano (2012) for developing countries. Similar results have been provided by Karbownik and Myck (2012), who find that among tertiary educated women in Poland, women with the strongest labour market attachment

select into the group with larger families (but a reverse can be observed among lower educated women). This suggests that in some countries or social groups, the selection mechanisms operate differently than in others. For example, the distinction between leisure-oriented and duties-oriented women may be stronger than the distinction between women that prefer pursuing paid jobs rather than getting involved in unpaid work at home. Also, as Matysiak (2011) emphasizes, in some countries, due to limited welfare state support for mothers, women may select into paid work in order to have (more) children. This strategy may provide women not only with financial security but also with legal entitlements to employment-based benefits that are guaranteed by the state only for the group of employees.

#### *DURATION OF THE EFFECTS*

The estimates obtained in “natural experiments” are not only much smaller than the ones presented in descriptive studies, but they also suggest that the negative effects of childbearing on female labour market career are only transitory. Specifically, Rosenzweig & Wolpin (1980) as well as Jacobsen, Pearce & Rosenbloom (1999) show that the effects of fertility on female labour supply are greatest directly after the child is born and then rapidly decline. The same can be concluded from Vere (2011), who shows that in 1990 a second child under 12 months old reduces labour force participation by 25%, but a child at the age of six diminished female labour supply by 13%, and by age 12 the effects of having children was no longer significantly different from employment among women with just one child.

The evidence on the changes of the effects of childbearing across the time since the birth comes only from US, and there are no studies that would show how this impact works under other institutional or cultural conditions. Nevertheless, the findings from the available studies suggest that on the macro-level, changes in fertility matter for the contemporaneous labour supply, but motherhood might have only limited impact on the labour market career of women if we look at labour market careers from the life course perspective.

## **V. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH**

The “natural experiments” related to human reproduction processes have opened up new opportunities with respect to causal inference regarding family size effects on female employment. Most studies carried out so far that implement such an approach, confirm that the number of children does have a negative effect on female employment, net of the impact of women’s preferences regarding involvement in home-based versus paid work. These studies provide also insight into mechanisms of selection of women into the group of mothers with larger families. In many countries, these are women with characteristics that decrease propensity for work, who opt for raising more children. However, it seems that this pattern of selection might not be universal. The evidence provided for some countries and social groups suggests that these are actually women with higher propensity for work that select into the group of women with more numerous offspring.

Studies based on “natural experiments” confirm the observations from earlier research showing that the effect of the number of children depends on parity and changes over time, as the child becomes older. Less attention has been paid to the variation in the magnitude of family size effect according to various living arrangements. With the exception of Bronars & Grogger (1994), no authors have explicitly elaborated on the differences in the (dis)employment effects experienced by women living in different family settings. And even the study by Bronars & Grogger (1994) compares only married and single mothers. Given the increasing diversity of family formation patterns in Europe and elsewhere, it would be interesting to see whether the impact of the number of children on female employment may differ depending on whether mothers can rely on their partners’ support or not, and whether formalising a union with a marital contract matters in this respect. Moreover, family members other than partners, such grandparents, can also provide help in taking care of children and thus relax the constraints of parents’ time resources (Aassve et al. 2011). Therefore, it could be of interest to see if their

involvement reduces the time conflict between work and childcare duties and decreases the negative effect of the number of children on maternal employment.

Apart from comparisons across diverse family settings, country comparisons could also provide interesting insights. One could suppose that depending on institutional and cultural arrangements, various countries or regions may differ in opportunities for combining paid work and family care. This suggests that a systematic comparison of the heterogeneity of family size effects in across societies could be a promising path for future research. Such an approach would acknowledge the postulate of Matysiak & Vignoli (2008) to consider family-related processes at the micro-level as influenced by macro-level factors.

## **ACKNOWLEDGEMENTS**

This research was carried out within a project "The impact of the number of children on the well-being of families in Europe" financed by a grant DEC-2011/03/D/HS4/04258 from the National Science Centre in Poland. I would like to express my gratitude for the support and hospitality that I received at Sociology Department at Umeå University, where the final version of this paper has been revised.

## **REFERENCES**

Aassve A., Arpino, B., & Goisis A. 2012. Grandparenting and mothers' labour force participation: A comparative analysis using the generations and gender survey. *Demographic Research*, 27(3), 53-84.

Agüero J. M. & Marks M. S. 2008. Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks. *American Economic Review- Papers and Proceedings* 98(2), 500-504.

Agüero J. M. & Marks M. S. 2011. Motherhood and Female Labor Supply in the Developing World: Evidence from Infertility Shocks. *Journal of Human Resources*, 46(4), 800-826.

Andersson G., Hank K., Rønsen M. & Vikat A. 2006. Gendering Family Composition: Sex Preferences For Children And Childbearing Behavior In The Nordic Countries. *Demography*, 43(2), 255-267.

Angrist J. D. & Pischke J.-S. 2010. The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics, *Journal of Economic Perspectives*, 24(2), 3-30.

Angrist J.D. & Evans W.N. 1998. Children and their parents' labour supply: evidence from exogenous variation in family size, *American Economic Review*, 88(3), 450-77.

Angrist J.D. & Evans W.N. 1996. Children and their parents' labour supply: evidence from exogenous variation in family size, National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5778.

Arpino B. & Aassve A. 2013. Estimating the causal effect of fertility on economic wellbeing: data requirements, identifying assumptions and estimation methods. *Empirical Economics*, 44(1), 355-385.

Blau F. D. & Kahn L. M. 2013. Female Labor Supply: Why Is the United States Falling Behind? *American Economic Review*, 103(3), 251-56.

Blundell R., Duncan A., & Meghir C. 1998. Estimating labor supply responses using tax reforms. *Econometrica*, 66(4), 827-861.

Bray J.W. 2005. Alcohol use, human capital, and wages. *Journal of Labor Economics*, 23(2): 279-312.

Brewster K. L. & Rindfuss R. R. 2000. Fertility and women's employment in industrialized nations. *Annual Review of Sociology*, 26, 271-296.

Browning, Martin 1992. Children and Household Economic Behavior *Journal of Economic Literature*, 30(3), 1434-75.

Cáceres-Delpiano J. 2012. Can We Still Learn Something From the Relationship Between Fertility and Mother's Employment? Evidence From Developing Countries. *Demography* 49(1):151-174

Caceres-Delpiano, J. 2006. The impacts of family size on investment in child quality, *Journal of Human Resources*, 41(4), 738-54.

Cain G. G. & Dooley M. D. 1976. Estimation of a model of labor supply, fertility, and wages of married women. *The Journal of Political Economy*, 84(4), S179-S199.

- Carrasco R. 2001. Binary choice with binary endogenous regressors in panel data. *Journal of Business & Economic Statistics*, 19(4), 385-394.
- Chun H. & Oh J. 2002. An instrumental variable estimate of the effect of fertility on the labour force participation of married women. *Applied Economics Letters*, 9(10), 631- 634.
- Cigno A. 1991. *Economics of the Family*, Oxford and New York: Clarendon Press and Oxford University Press.
- Cristia J. P. 2008. The Effect of a First Child on Female Labor Supply Evidence from Women Seeking Fertility Services. *Journal of Human Resources*, 43(3), 487-510.
- Cruces G. & Galiani S. 2007. Fertility and female labor supply in Latin America: New causal evidence. *Labor Economics*, 14(3), 565-573.
- Daouli J., Demoussis M. & Giannakopoulos N. 2009. Sibling-sex composition and its effects on fertility and labor supply of Greek mothers. *Economics Letters*, 102(3), 189-191.
- Del Boca D. & Locatelli M. 2002. Fertility and Women's Work Status. Report No 6 for the project "The Rationale of Motherhood Choices: Influence of Employment Conditions and of Public Policies".
- Duncan G. J., Magnuson K. A. & Jens L. 2004. The Endogeneity Problem in Developmental Studies. *Research in Human Development*, 1(1-2), 59-80
- Engelhardt H., Kohler H.-P., Prskawetz A. 2009. *Causal Analysis in Population Studies*, The Springer Series on Demographic Methods and Population Analysis, 23, 1-7.
- Ermisch J. F. 1989. Purchased child care, optimal family size and mother's employment Theory and econometric analysis. *Journal of Population Economics*, 2(2), 79-102.
- Foley M. C. & York G. A. 2005. The Effect of Children on Female Labour Supply in the United States From 1950 to 2000. *Mimeo*.
- Francesconi M. 2002. A joint dynamic model of fertility and work of married women. *Journal of Labor Economics*, 20(2), 336-380.
- Gangadharan J. & Rosenbloom J. L. 1996. The Effects of Child-bearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment. Working Paper 5647, National Bureau of Economic Research, Cambridge, MA.

Haan P. & Wrohlich K. 2011. Can child care policy encourage employment and fertility?: Evidence from a structural model. *Labour Economics*, 18(4), 498-512.

Habbema J.D.F., J.Collins, H.Leridon, J.L.H.Evers, B.Lunenfeld & E.R. Velde 2004. Towards less confusing terminology in reproductive medicine: a proposal. *Human Reproduction*, 19(7), 1497-1501.

Heckman J. & R. Willis 1976. 'Estimation of a Stochastic Model of Reproduction: An Econometric Approach' in Terleckyj N. (ed.) *Household Production and Consumption*, New York: Columbia University Press.

Heckman J. & Urzua S. 2010. Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156(1), 27-37.

Heckman J. 1976. A Life Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy*, 84(4), S11-S44.

Hotz V. J. & Miller R. A. 1988. An Empirical Analysis of Life Cycle Fertility and Female Labor Supply." *Econometrica*, 56(1), 91-118.

Hotz V. J., McElroy S. W. & Anders S. G. 2005. Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment. *The Journal of Human Resources*, 40(3), 683-715

Hotz, V.J., Mullin C. & Sanders S. 1997. Bounding Causal Effects Using Data From a Contaminated Natural Experiment: Analysing the Effect of Teenage Childbearing. *Review of Economic Studies*, 64(4):575-603.

Iacovou M. 2001. Fertility And Female Labour Supply. ISER Working Papers Number 2001-19.

Jacobsen J. P. & Pearce J. W. & Rosenbloom J. L. 1999. The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment. *Journal of Human Resources*, 34(3), 449-474.

Javornik J. 2012. Rethinking comparative childcare policy analysis: Example of Central and Eastern Europe. Working paper 10/2012, Department of Sociology, Umeå University.

Johansson E, Alho H, Kiiskinen U., Poikolainen K. 2007. The association of alcohol dependency with employment probability: evidence from the population survey 'Health 2000 in Finland'. *Health Economics*, 16(7), 739-754.

Joshi, H. 1998. The opportunity costs of childbearing: More than mothers' business. *Journal of Population Economics*, 11(2), 161-183.

Karbownik K. & Myck M. 2012. For Some Mothers More Than Others: How Children Matter for Labour Market Outcomes When Both Fertility and Female Employment Are Low. IZA DP No. 6933

Kline, J., Stein, Z. A., Susser, M., & Warburton, D. 1977. Smoking: a risk factor for spontaneous abortion. *New England Journal of Medicine*, 297(15), 793-796.

Lehrer E. & Nerlove M. 1986. Female Labor Force Behavior and Fertility in the United States. *Annual Review of Sociology*, 12, 181-204.

Lehrer E., Grossbard-Shechtman Sh. & Leasure, J. W. 1996. Comment on 'A Theory of the Value of Children. *Demography*, 33(1), 133-136.

Levine, P. B., Gustafson, T. A. & Velenchik, A. D. (1997). More bad news for smokers? The effects of cigarette smoking on wages. *Industrial and Labor Relations Review* 50(3), 493-509.

Martin J.A. & Park M.M. 1999. Trends in twin and triplet births: 1980-97. *National Vital Statistics Reports* 47:1-17.

Matysiak A. & Vignoli D. 2008. Fertility and women's employment: a meta-analysis. *European Journal of Population*, 24(4), 363-384.

Matysiak A. & Węziak-Białowolska D. 2013. Country-specific conditions for work and family reconciliation: an attempt at quantification, Institute of Statistics and Demography Working Paper No 31.

Matysiak, A. 2011. Interdependencies between fertility and women's labour supply Vol. 17. Springer.

Michalopoulos C., Robins P. K. & Garfinkel I. 1992. A structural model of labor supply and child care demand. *Journal of Human Resources*, 27(1), 166-203.

Mincer J. & H. Ofek 1982. Interrupted Work Careers. *Journal of Human Resources* 17(1), 3-24.

Moffitt R. 2005. Remarks On The Analysis Of Causal Relationships In Population Research, *Demography*, 42(1), 91-108.

Moffitt R. 2003. Causal Analysis in Population Research: An Economist's Perspective," *Population and Development Review*, 29(3), 448-458.

- Moffitt R. A. 1984. Profiles of Fertility, Labour Supply, and Wages of Married Women: A Complete Life-Cycle Model. *Review of Economic Studies*, 51 (2), 63-78.
- Nam K. 2010. The Effect of Having More Children on Women's Labour Force Participation in Korea: An Analysis Using Instrument Variables. *Labour*, 24(3), 333-356.
- Namboodirni N. 1972. Some Observations on the Economics Framework for Fertility Analysis. *Population Studies*, 26, 185-206.
- Pebley A. R., & Westoff C. F. 1982. Women's sex preferences in the United States: 1970 to 1975. *Demography*, 19(2), 177-189.
- Rosenzweig M. R. & Wolpin K. 1980. Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models, *Journal of Political Economy* 88(2), 328-348.
- Rosenzweig M. R. & Wolpin K. I. 2000. Natural 'natural experiments' in economics. *Journal of Economic Literature*, 38(4), 827-874.
- Rosenzweig M. R. & Zhang, J. 2009. Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy. *The Review of Economic Studies*, 76(3), 1149-1174.
- Schultz T. P. 1980. Estimating Labor Supply Functions for Married Women. In: Smith J. *Female Labor Supply*. Princeton, New York: Princeton University Press.
- Schultz T. P. 2008. Population Policies, Fertility, Women's Human Capital, and Child Quality, *Handbook of Development Economics*, Elsevier.
- Vere J. 2011. Fertility and parents' labour supply: new evidence from US census data. *Oxford Economic Papers*, 63 (2), 211-231.
- Weiss Y. & Gronau R. 1981. Expected Interruptions in Labour Force Participation and Sex-Related Differences in Earnings Growth. *Review of Economic Studies*, 48(4): 607-619.
- Weiss Y. 1972. On the Optimal Lifetime Pattern of Labour Supply. *Economic Journal*, 82(382), 1295-1315.
- Willis, Robert J. 1973. A New Approach to the Economic Theory of Fertility Behavior. *Journal of Public Economics*, 81(2), S14-S64.
- Wrohlich K. 2011. Labor supply and child care choices in a rationed child care market, Discussion Papers No. 1169, German Institute for Economic Research, DIW Berlin.

Xia, L. 2010. How Children Affect Mother's Labor Market Outcomes: Evidence Based on Pregnancy Failures. Mimeo.