IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Essays on schooling, gender, and parental leave

Elly-Ann Johansson

DISSERTATION SERIES 2010:1
Presented at the Department of Economics, Uppsala University

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

This doctoral dissertation was defended for the degree of Doctor in Philosophy at the Department of Economics, Uppsala University, January 29, 2010. Essay 1 is a revised version of IFAU Working paper 2008:21 and essay 4 is also published as IFAU Working paper 2010:4.

ISSN 1651-4149

Doctoral dissertation presented to the Faculty of Social Sciences 2010

Abstract

Dissertation presented at Uppsala university to be publicly examined in Hörsal 2, Ekonomikum, Kyrkogårdsgatan 10, Uppsala, Friday, January 29, at 10.15, for the degree of Doctor of Philosophy. The examination will be conducted in English. JOHANSSON, Elly-Ann, 2009, Essays on schooling, gender, and parental leave; Department of Economics, Uppsala University, Economic studies 121, 131 pp, ISBN 978-91-85519-28-6, ISSN 0283-7668
urn:nbn:se:uu:diva-111210 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-111210)

This thesis consists of four self-contained essays.

Essay 1: Mixed-aged classes (MA-classes) are a common phenomenon around the world. In Sweden, these types of classes increased rapidly during the 1980:s and 1990:s, despite the fact that existing empirical support for MA-classes is weak. In this paper, we estimate the effect of attending an MA-class during grades 4-6 on students' cognitive skills. Using a unique survey with information on students, parents and teachers, we are able to control for many factors that could otherwise bias the results. We find a negative effect on short-run cognitive skills, as measured by grade 6 cognitive tests, and this effect is robust to a rigorous sensitivity analysis.

Essay 2: We examine whether the impact of pre-school interventions on cognitive skills differs by immigrant background. The analysis is based on Swedish data containing information on childcare attendance, rich family background information, the performance on cognitive tests at age 13, and long-run educational attainment for cohorts born between 1967 and 1982. We find that childcare attendance reduces the gap in language skills between children from immigrant backgrounds relative to native-born children. We find no differential effects on inductive skills, however. Nor does childcare appear to affect the distribution of long-run educational attainment.

Essay 3: This paper estimates the effect of child gender on mothers' and fathers' parental leave. The focus on Sweden, a highly gender equal society, yields additional knowledge on the prevalence of gender biases in industrialized countries. The results show that a first born son increases fathers' parental leave with 0.6 days (1.5 percent) and decreases mothers' leave by a similar amount, leaving the total leave unchanged. Both the sign and size of this effect is in line with previous research. However, there are interesting differences between groups that departs from previous studies. Non-traditional families, with high maternal relative earnings and/or

educational levels, show even larger gender biases, indicating that it may be mothers, rather than fathers, that are the driving force behind this child gender bias.

Essay 4: This paper investigates the effect of parental leave – both own and spousal – on subsequent earnings using different sources of variation. Using fixed-effects models, and in line with previous results, parental leave is found to decrease each parent's future earnings. Also spousal leave is important, but only for mothers. In fact, each month the father stays on parental leave has a larger positive effect on maternal earnings than a similar reduction in the mother's own leave. Using two reforms of the parental leave system as exogenous sources of variation yields only imprecisely estimated effects, even though the reforms had a strong effect on parental leave usage. However, the point estimates tentatively suggest effects in the same range or larger than the fixed-effects model found.

Acknowledgements

As always, this thesis had not been possible without the contributions from several people. First and foremost, I am very grateful to my supervisors Peter Fredriksson and Per Johansson. Peter for careful reading of every draft and for the ability to explain difficult issues in an easy-to-understand way, Per for enthusiasm and valuable econometric expertise, and both of you for patience with my little detours along the way. I would also like to thank Per-Anders Edin for introducing me to interesting research contacts and my colleagues Erica Lindahl and Caroline Hall for nice co-work. My opponents at the final and licentiate seminars, Jonas Lagerström and Peter Skogman Thoursie, have made useful suggestions. Björn Öckert has given important help, especially with the last two papers in the thesis. Numerous other people – colleagues, the administrative staff, and seminar participants – have also contributed in different ways.

Equally important is the social environment. In particular, I want to thank my first-year class mates Jenny, Micke, Matilda, Håkan, Louise and Xiao: Jenny for (too much?) fun, both at and after work, Micke for being such a good friend (and also an ever-lasting source of departmental knowledge), and all of you for interesting discussions and nice coffee-breaks. Thanks also to my "Berkeley-family" for a funny and (mostly) sunny semester abroad. The rest of you – no one mentioned and none forgotten – have also made the time at the department a very nice experience.

Needless to say, thanks also to family and friends. And Micke and Elmer.

Elly-Ann Johansson, December 2009, Uppsala

Contents

Introduction

What constitutes good research? Most importantly, of course, it must address important issues. Second, it must be able to deliver reliable answers to the questions asked. This thesis consists of four essays on quite different topics, all with different policy relevance and different strategies to (reliably) identify the effect of interest.

The first two papers deal with the effects of mixed-age classes (Essay 1) and pre-schools (Essay 2) on children's cognitive achievement. These issues are certainly important from a policy perspective. In Sweden, mixed-classes increased rapidly during the 1980s and 1990s, and almost 40 percent of all students in grade 3 attended an MA-class during 2002-2008 (calculations using Statistics Sweden's database). This was partly a response to a growing belief in the pedagogical benefits of these types of classes, but the scientific support for such a belief was, and still is, poor. For example, most previous studies are based on small and often non representative samples (Veenman, 1995). Even more widespread is the use of public day care – over 80 percent of all one to five year olds were enrolled in a preschool in 2005, compared to less than 10 percent only some 30 years earlier (calculations using databases from The Swedish National Agency for Education).

Yet, despite numerous studies in both areas there is still no consensus among researchers (Little, 2001; Mason and Burns, 1996; Veenman, 1995; Swedish National Institute of Public Health, 2009; Waldfogel, 2002). Are pre-schools good or are the kids better off at home? And what are the distributional effects – do pre-schools increase or reduce inequality? Do mixed-age classes, where children of different age and school experience are mixed, constitute a better or worse learning environment than ordinary classes?

The reason for the lack of consensus is that in the social sciences in general, and perhaps in fields like child care and schooling in particular, it is difficult or impossible to perform randomized experiments. Instead, observational data are available, but these potentially yield misleading answers. For example, two stylized facts are that well educated parents in Sweden place their children in preschool more often than other parents, and that children from well-educated homes perform on average better in school. If we observe that children who attended preschool at young ages perform better in school later on compared to other children, it could be a causal effect – preschools cause children to perform better – but it could also be merely a correlation (highly educated parents put children in child care *and* have high performing children) in which case we have learned nothing about the effect of preschools. But, by adding study after study and using continuously better and different types of empirical strategies, research may help us get closer to identifying the causal effect of interest.

One solution to the identification problem is to "control" for important observable characteristics in a regression adjustment setting, hoping that is enough to rule out confounding factors. By different means it is also possible to evaluate the sensitivity of the results with respect to the included control variables, and thereby gain additional insight into whether the estimates are causal or merely correlations. This is the strategy used in the first and second paper.

Another solution to the identification problem is to find changes in the environment that mimics a randomized experiment. Such changes are difficult to find, and to the extent that the changes are truly exogenous, this strategy is often considered more reliable in terms of identification. The third and fourth paper, both dealing with gender issues in relation to parental leave, employ such plausibly exogenous variation.

Essay 3 examines how child gender affects parental leave. It utilizes the gender of the first child as exogenous variation (the gender of later-born siblings has been shown endogenous). While this paper has the perhaps cleanest identification, it is at the same time of limited policy relevance (at least ex post, given the findings). Its main contribution is more interdisciplinary, as it questions the use of child gender as exogenous instruments in other applications. For example Bennedsen et al, 2007, investigate the effect of family-CEO:s on firm performance. They instrument the choice of CEO – external or from the family – by the gender of the first-born child. The idea is that firm control more often is passed onto sons. Their strategy may be questioned if child gender affects parental behavior in ways that in turn could affect firm performance.

The third paper also sheds more light on the prevalence of gender biases in the industrialized world. While it for long was believed that such biases existed only in the developing world (as indicated by "missing girls" and skewed sex-ratios), new evidence shows that the same types of biases exist also in industrialized countries; the gender biases in the industralized countries are expressed differently, however, for example in terms of higher marital happiness and stability in families with sons or different fertility patterns depending on child gender (Lundberg, 2005).

Essay 4 is more policy oriented, as it investigates how a more equally shared parental leave affects future earnings. The question is important, not least in light of the developments on the labor market. While there has been dramatic rises in female labor force participation and hours worked during the last decades, women still take the lion's share of housework and child care. For example, Figure 1 shows the development of parental leave usage for Swedish mothers and fathers over time. While most mothers use the vast majority of parental leave – around 400–450 days – fathers' share of parental leave is still small. During most part the 1990s, fathers' parental leave averaged around 40–50 days and the mean number of days was still only around 80 in the beginning of 2000. These clear differences in child care responsibilities are sometimes suggested as one potential part of the explanation for the remaining, unexplained female-to-male earnings gap (see for example Lundberg and Pollak, 2007). If so, increasing the fathers' share of parental leave would be one way of closing the gender gap in earnings.

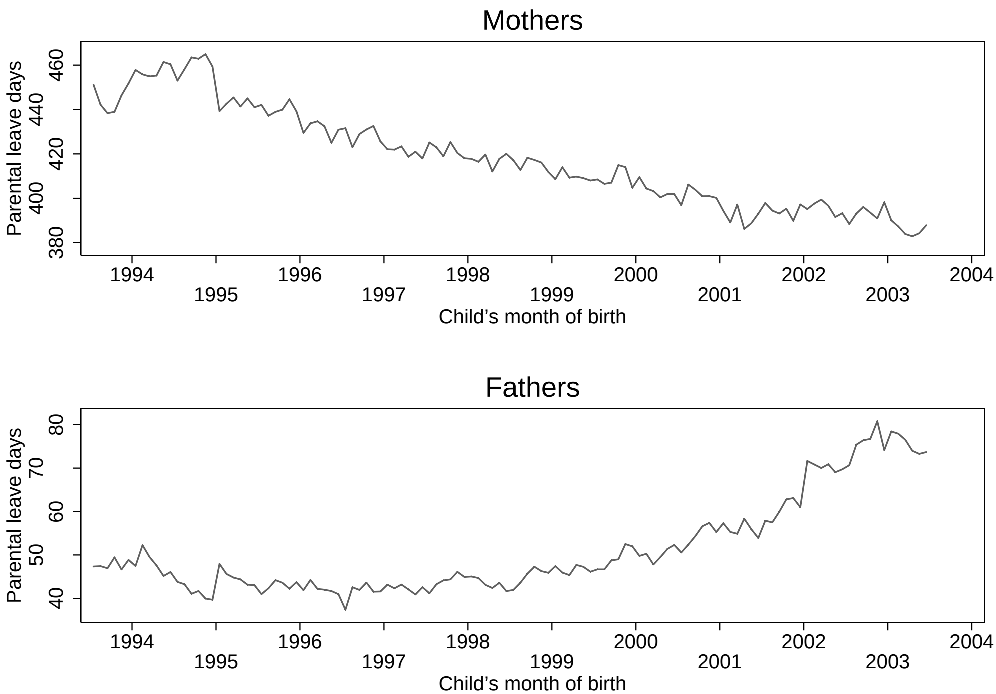

Figure 1 Mothers' and fathers' parental leave days over time
Notes: In this figure, parental leave is calculated up to child age 3 for the first-born child.

Estimating the effect of a more equally shared parental leave on earnings is, however, difficult since families where the father uses relatively large amounts of parental leave may differ in important aspects from other types of families. Therefore, in addition to ordinary regression adjustment, the paper also utilizes two reforms of the parental leave system as exogenous variation in parental leave.

Below, I give a brief summary of each paper and its main findings.

Essay 1: The effect of MA-classes in Sweden

The first essay, "The effect of mixed-age classes in Sweden", investigates the effect of mixed-age classes (henceforth MA-classes) on students' cognitive skills. In MA-classes children of different age and school experience are mixed into one single class, sometimes out of demographic necessity, sometimes due to a belief that these classes have pedagogical benefits.

Our data set is a stratified sample of 4584 students who attended grade 3 in 1992, which means that most students were born in 1982 and finished 9th grade in 1998. We compare children in MA-classes with children in traditional classes, and given an unusually rich data set with information on both students, parents, teachers and schools, we hope that controlling for these observable factors will be enough to identify the causal effect of MA-classes on cognitive outcomes. For example, we have information on whether the family has made an active school choice or simply accepted the school located closest to home. This variable is potentially an important control variable for selection of students into class types.

The results show that children in MA-classes score lower on cognitive tests in grade 6 compared to children from traditional classes, but we cannot detect any longer run effects on grade 9 credits. These results are also robust to a number of different sensitivity analyses, including adding a control for ability.

Essay 2: Do pre-school interventions further the integration of immigrants? Evidence from Sweden

The second essay, "Do pre-school interventions further the integration of immigrants? Evidence from Sweden" studies the distributional effects of pre-schools. The prime focus is on whether the effects of pre-schools vary by immigrant status. The data contain information on roughly 10 percent of all families with children from four different birth cohorts (born in 1967, 1972, 1977 and 1982, respectively). During this period, there was a dramatic rise in female labor force participation, which resulted in increased child care attendance and a subsequent change in the composition of pre-school children – they were drawn relatively more from the upper end of the income distribution.

The main dependent variables are results from two different cognitive tests in grade 6, one verbal test and one numerical test. The main independent variable is an indicator variable for child care attendance. Regression adjustment is used to identify causal effects, and different alternative specifications are used to investigate how sensitive the results are to the included control variables. These sensitivity analyses show that we cannot credibly identify the mean effect of child care; however, the distributional effects are robust to alternative sets of control variables.

The results show that pre-schools seem to reduce inequality, at least along certain dimensions: the difference in verbal skills between native and immigrant children is significantly reduced among children who attended a pre-school. However, there are no effects on inductive skills or on the distribution of long run educational attainment.

Essay 3: Gender bias in a gender equal society – evidence from Swedish parental leave use

The third essay, "Gender bias in a gender equal society – evidence from Swedish parental leave use", asks whether parents treat their children differently depending on their child's gender. The data used is a population wide register panel data set of native Swedish families whose first child was born during 1993-2005. There are virtually no missing variables or attrition.

The results show that Swedish parents are not gender neutral. Fathers take an additional 0.6 parental leave days (1.5 percent) for sons; mothers conversely reduce parental leave by around 0.6 days, leaving the total parental leave unchanged. From a policy perspective, the effects are very small in magnitude and unlikely to have any effect on child development. However, they do show that also parents in Sweden – the most gender equal country in the world according to United Nations rankings (United Nations, 1995) – are not neutral with respect to child gender. Instead, both the sign and size of the effects are in line with studies from other industrialized countries on how child gender affects fertility and marital status (Dahl and Moretti, 2008).

In addition, and in clear contrast to studies from other countries (see for example Dahl and Moretti, 2008), it seems to be the mothers who are the driving force behind these gender biases. Assuming that the relative earnings and/or educational level of the mother is a relevant proxy for her bargaining power within the family, and assuming also that a higher bargaining power implies higher influence over the parental leave decision, we can shed light on this issue by estimating the model separately for subgroups with different maternal relative educational/earnings levels. Perhaps suprisingly, the general pattern is that the gender bias is larger, the more bargaining power the mother has.

Finally, the paper also investigates how child gender affects temporary parental leave (parental leave used for taking care of older children when they are sick). (To that end, Linda data is used, which is a random and representative sample of around three percent of the Swedish population and their family members.) Again, fathers are home more with sons than with daughters and for the temporary parental leave the gender effect is even larger – a son increases fathers' leave by 3.3 percent.

Essay 4: The effect of own and spousal parental leave on earnings

The fourth essay, "The effect of own and spousal parental leave on earnings" investigates the effect of parental leave on earnings. It fits into a broader literature on how career interruptions in general and parental leave in particular affect subsequent earnings through different channels – loss of human capital during the leave, signaling of work commitment and/or statistical discrimination (Albrecht et al, 1999; Datta Gupta and Smith, 2002; Gangl and Ziefle, 2009; Görlich and De Grip, 2009; Mincer, 1974; Mincer and Polachek, 1974; Mincer and Ofek, 1982; Ruhm, 1998; Skyt Nielsen, 2009).

But it departs from previous studies in two main ways. First, it also considers spousal parental leave, an issue mostly ignored in previous work. The hypothesis is that if fathers take more parental leave, they might "learn" to take a larger share of housework and child care also in the future, which in turn may feed back onto female labor market behavior. Second, it utilizes several sources of variation to identify causal effects. In addition to cross-sectional and fixed-effects models, it uses two reforms of the parental leave system as potentially exogenous variation in parental leave and estimates difference-in-differences (DD) or triple differences (DDD) models.

The results from fixed-effects models show that, in line with previous research, own parental leave lowers future earnings. However, and more interesting, also spousal parental leave is important, but only for mothers. Each month that the father stays on parental leave increases the mother's future earnings by almost 7 percent, and this effect is even larger than a similar reduction in the mother's own leave. This suggests that spousal (lack of) involvement in child care and parental leave could explain part of the remaining earnings gap.

This interpretation of the results rests on the fixed-effects assumption of no time-variant unobserved heterogeneity. In particular, parental leave should not respond to income shocks. Using the reforms as exogenous variation is one way of relaxing this assumption (the DDD model instead rests on the assumption that reform exposure is exogenous, which in practice means that previous income patterns should not affect whether families give birth before or after the reform cutoff).

Unfortunately, using the reforms as exogenous variation produces only imprecisely estimated effects (despite a strong effect on parental leave use), but the point estimates tentatively suggests effects in the same range or larger than found in the fixed-effects specification.

References

Albrecht, J.W., P-A Edin, M. Sundström and S. B. Vroman (1999): Career interruptions and subsequent earnings: a reexamination using Swedish data, The Journal of Human Resources, vol. 34, no. 2

Bennedsen, M., K. Meisner Nielsen, F. Perez-Gonzalez and D. Wolfenzon (2007), Inside the family firm: the role of families in succession decisions and performance, Quarterly Journal of Economics, vol. 122, no 2

Dahl, G. B., and E. Moretti (2008), The demand for sons, Review of Economic Studies, vol. 75, issue 4

Datta Gupta, N. and N. Smith (2002): Children and career interruptions: the family gap in Denmark, Economica, vol. 69

Gangl, M. and A. Ziefle (2009): Motherhood, labor force behavior, and womens careers: an empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States, Demography, vol. 46, no. 2

Görlich, D. and A. De Grip (2009): Human capital depreciation during hometime, Oxford Economic Papers, vol. 61, supplementary issue 1

Little, A.W. (2001): Multigrade teaching: towards an international research and policy agenda, International Journal of Educational Development, vol. 21, no. 6

Lundberg, S. and R.A. Pollak (2007): The American family and family economics, NBER working paper no. 12908, National Bureau of Economic Research

Lundberg, S. (2005), Sons, daughters, and parental behaviour, Oxford Review of Economic Policy, vol. 21, no. 3

Mason, A. and R.B. Burns (1996): ‘Simply no worse and simply no better’ may simply be wrong: a critique of Veenman’s conclusion about multigrade classes, Review of Educational research, vol. 66, no. 3

Mincer, J. and H. Ofek (1982): Interrupted work careers: depreciation and restoration of human capital, The Journal of Human Resources, vol. 17, no. 1

Mincer, J. and S. Polachek (1974): Family investments in human capital: earnings of women, The Journal of Political Economy, vol. 82, no. 2

Mincer, J (1974): Schooling, experience, and earnings, New York: Columbia University Press

Ruhm, C. (1998): The economic consequences of parental leave mandates: lessons from Europe, The Quarterly Journal of Economics, vol. 113, no. 1

Skyt Nielsen, H. (2009): Causes and consequences of father’s child leave: evidence from a reform of leave schemes, IZA discussion paper no. 4267, Institute for the Study of Labor, Bonn, Germany

Swedish National Institute of Public Health (2009): Child day care center or home care for children 12-40 months of age – what is the best for the child? A systematic literature review, report no. 2009:09

United Nations (1995): Human development report 1995: gender and human development, Oxford University Press

Veenman, S. (1995): Cognitive and non cognitive effects of multigrade and multi-age classes: a best evidence synthesis, Review of Educational Research, vol. 65, no. 4

Waldfogel, J. (2002): Child care, women's employment, and child outcomes, Journal of Population Economics, vol. 15, no. 3

Essay 1: The effect of mixed-age classes in Sweden[1]

(together with Erica Lindahl)

1 Introduction

Mixed-age classes are a common phenomenon in schools both in Sweden and in other countries. In mixed-age classes (henceforth MA-classes), students from different grades are mixed into one class for two major reasons: either out of demographic and economic necessity (too few children in each grade to form a class) or because it is believed that these classes have pedagogical benefits. For example, it is argued that students of different age and school experience interact and learn from each other. This belief contributed to the rapid increase of MA-classes in Swedish schools during the 1980:s and 1990:s.

However, the scientific evidence on the effects of mixed-age grouping is ambiguous. Among the Swedish studies used to motivate the introduction of MA-classes there is, to our knowledge, no study using representative samples. Many studies are simply questionnaires collected among teachers in MA-classes (Andrae Thelin, 1991; Edlund and Sundell, 1999). International studies are available, but they are of varying quality (with very few studies using representative samples) and yield contradictory results. However, most studies conclude that the effect, if any, is small in magnitude.

From an economic point of view, investigating the effect of MA-classes is important since it may be one possible way towards greater cost-efficiency within schools. If it is the case that MA-classes, as is often claimed, are a less expensive way to organize students than traditional classes, and if the students in these classes perform equally well or better than students in traditional classes, introducing MA-classes in a larger scale would be an

[1] We are grateful to Peter Fredriksson and Per Johansson for valuable guidance. We would also like to thank Mikael Elinder, Patrik Hesselius, Jenny Nykvist, Peter Skogman Thoursie, Andreas Westermark and seminar participants at the Department of Economics, Uppsala University, and an anonymous referee, for valuable suggestions and comments. Åsa Arnell is acknowledged for the research idea. The financial support from the Swedish Council for Working Life and Social Research, FAS (dnr 2004-1222) is also acknowledged.

efficient way towards reduced costs and/or increased student performance. This is particularly interesting in relation to one of the most debated policies in the economic and educational literature during the last years, namely reducing class-size. In contrast to class-size reductions, introducing MA would imply practically no extra costs.[2]

Examining the effect of MA-classes also sheds light on the question of how knowledge is produced. Economic research has mainly focused on quantitative aspects of education – if and how much resources matter for student achievement. But equally important are more qualitative aspects of the educational production function, and the effect of MA-classes is one such aspect.

The purpose of this paper is to estimate the effect of MA-classes in Sweden on students' cognitive skills. We focus both on short-term effects on grade 6 cognitive tests and on long run effects on grade 9 credits. We also allow the effect of attending an MA-class to vary between different groups of students considered potentially important: girls, low performing students and students with non-Swedish background.

The analysis is based on a rich and representative data set. In addition to register data on important socioeconomic variables, we have access to a unique survey with information on parents and teachers and their attitudes towards school related issues. These data allow us to control for many potential selection problems and perform a rigorous sensitivity analysis.

The results show a negative effect of attending an MA-class in grades 4–6 on the grade 6 cognitive tests. This effect is not statistically different for girls, low performing students or students with a non-Swedish background. The point estimate of the effect of MA-class-attendance on grade 9 credits is negative but not statistically significant.

2 Background

In this section we discuss the concept of MA and the prevalence of these types of classes across time and countries. We also present arguments used for and against these classes and review empirical evidence of the effects of MA-classes on child outcomes.

2.1 MA-classes: development and definition

MA-classes arise for two main reasons: either through economic necessity (too few children in an area to form a class or too few teachers to cover all grades) or through choice (a belief that MA-classes are pedagogically superior). Generally, the knowledge is poor about the prevalence of MA-

[2] For reviews of the class-size literature, see Krueger (2003) and Hanushek (1999).

classes across time and countries (Little, 2001) but some general patterns can be described. Historically, MA-classes were the only possible way of organizing schools due to low population density. Still today, MA-classes are mainly found in rural areas where they have been formed out of necessity (Little, 2001). However, from the 1960s and onwards, a belief in the pedagogical benefits of MA-classes started to spread in many countries and today, MA-classes within urban areas is usually the result of an active pedagogical choice (Little, 2001).

In Sweden, the number of MA-classes increased rapidly from 1980 and onwards (Vinterek 2001; 2003). In 2000, approximately one third of all Swedish students in the first three years of school attended MA-classes and about one fourth of the students in grades 4 and 5. That is nearly twice as many as only five years earlier. The share of students attending an MA-class during the last three years of compulsory schooling in Sweden is still rather small; about 2 percent of all students in these grades were in mixed-age groups between 1996 and 1998.

We do not know whether this rapid increase in the number of MA-classes in Sweden is due to pedagogical reasons or economic reasons (Vinterek, 2003). There is some evidence that pedagogical motives dominated in the lower grades (1-3) whereas economic motives dominated in the higher grades (4-6) (Sandquist, 1994). In grades 7–9, mixed-age classes are scarce, and if they do exist, they tend to be motivated by demographic necessities (Sandquist, 1994).

The initiative to start an MA-class has usually come from groups of teachers within a school, often supported by the school management (Vinterek, 2003). However, since the beginning of 1990 it seems to be the case that MA-classes have been introduced by politicians against the will of teachers and parents (Vinterek, 2003; Edlund and Sundell, 1999; Sundell, 2002 and Sandquist, 1994). There is also evidence that mixed-age classes are more prevalent in schools with many low performing students (Vinterek, 2003).

The basic definition of a mixed-age class is a class consisting of students of different age and from different grades (as compared to a conventional class where all students are from the same grade and all or most students are of the same age). In practice, the term mixed-age education also often implies a different type of teaching method, although there is little consensus about what characterizes this teaching method.

There are a number of different teaching strategies available: teaching the whole class simultaneously, ability grouping within the class irrespective of grades and grouping by grades in some subjects while teaching the whole class simultaneously in others. Which strategy is most prevalent is unknown and according to Little (2001), mixed-age teaching is "invisible" in textbooks, syllabi and teachers' education. For the Swedish setting, there is some evidence suggesting that students in MA-classes work more individually (Sandqvist, 1994; Vinterek, 2003)[3]. There are also tendencies to grade-specific teaching. However, there are large differences depending on subject. Social sciences are often taught to all grades simultaneously, and leave large possibilities for group work and a thematic organization of the subject. In subjects like Mathematics and Athletics, teaching is more often done separately for each grade. This can be achieved in different ways. Sometimes one grade within the class works individually with one subject while the other grade(s) listens to the teacher lecturing. In other cases, the highest graders stay in school later in the afternoon and have time to learn more advanced Mathematics when their younger classmates have left for the day (Sandqvist, 1994; Vinterek, 2003).

2.2 Arguments for and against MA-classes

Mixed-age classes may differ from conventional classes in two ways: the composition of students and the teaching methods. Most arguments in favour of MA-classes focus on the former, i.e. the effects of greater student heterogeneity. In the following we give an exposition of the most commonly used arguments for MA-classes. Since the literature is mainly concerned with the supposed benefits of MA-classes also this exposition will be one-sided; this does not mean that the pro-arguments have more empirical support.

Veenman (1995) discusses the following benefits of MA-classes: MA-classes are claimed to enhance the children's security and confidence as they form relationships with a wider variety of children. MA-classes also invite cooperation, and children benefit from learning from and teaching each other. Furthermore, MA-classes are considered to have a more relaxed atmosphere, and to be more stimulating as children from different ability levels meet. It is also claimed that the self-concepts of slower, older students are specially enhanced when they are asked to tutor younger students.

In order to motivate the introduction of MA-classes in Sweden, the following arguments have been used by many local politicians in local school directives (Sandqvist, 1994). MA-classes enable greater adaptation to individual maturity in different subjects and generate greater social training since the group is more heterogeneous with respect to age. In addition, mixed age grouping is claimed to give rise to more acceptance for deviating behavior among classmates.

[3] One reason could be the large heterogeneity within the class, making cooperation between students and group activities more difficult since they are at different knowledge levels. This implies that learning takes place through quiet reading and writing more than through listening and speaking. This is somewhat contradictory; since one common argument for MA-classes is that the larger heterogeneity within the class enhances learning through group activities.

Some local politicians also refer to the pedagogical idea that students are assumed to be naturally curious and hungry for knowledge and that children spontaneously learn from each other and willingly teach each other.[4] Given this view of schooling and children, a more heterogeneous group is desirable. Another argument, connected to the former, is that the new post modern information intensive society requires knowledge about how to *search* for information. To work in project teams and to cooperate among students in order to search for information are new features in the school directives that is claimed to fit well with MA-teaching.

Sundell (1995) also describes the arguments used in directives from the former Swedish National Agency for Education (Skolöverstyrelsen). Among the arguments in favour of MA-classes, the supposed positive impact on students' cognitive development is claimed to stem from the teaching adapted to the individual that is connected with MA-classes, as well as the idea that younger students learn from their older peers. The reason for the former argument is that in an MA-class, working groups are formed in accordance with the individual child's mental maturity rather than its actual age.

Further, it is often claimed that the individually adapted teaching connected with MA-classes specially benefit low performing students. The reasons are several. First, it is argued that the individually adapted teaching results in more teaching time to those in special need. Second, teaching in an MA-class is to a higher degree organized in small groups, which benefit low performing students. Finally, as stated above, in an MA-class low performing students have the possibility to compare themselves with younger children and in this way they do not need to perform worst.

Arguments are sometimes contradictory. For example, student heterogeneity are viewed as beneficial either because heterogeneity in itself is positive or because this heterogeneity allows more ability grouping, i.e. less heterogeneity within the classroom. As another example, while some claim that MA-classes give the teachers a better working environment as only a share of the class is new every year (Sundell, 1995), others instead argue that MA-classes impose a greater workload on the teachers and that most teachers are not adequately prepared to deal with MA-groups (Veenman, 1995).

In sum, there is no theoretical consensus about the mechanisms behind MA-classes.

2.3 Empirical studies

The empirical evidence on the effects of MA-classes is ambiguous and many studies are of poor quality. For example, Veenman (1995) summarizes

[4] The Montessori pedagogy is mentioned in some local school directives (Sandqvist, 1994).

evidence from 56 international studies that investigate the effect of MA-classes on both cognitive (test scores and grades) and non-cognitive (for example self-concept, adjustment and attitudes towards school) child outcomes. There were no experimental studies at all, and virtually no studies based on representative samples of the student population with well-defined treatment- and comparison groups. Many studies did not even make any attempts to condition on initial differences between students in MA-classes versus traditional classes.

The studies yield contradictory evidence, and when summarizing the results from the studies of best quality, the average effect of attending an MA-class becomes zero. The reason for this zero effect is discussed by Mason and Burns (1996). They argue that selection of better students and/or teachers into MA-classes are counteracted by less effective instruction in these classes. In another review (Mason and Burns, 1997) they investigate research on ”combination classes” – mixed-age classes formed out of economic neccessity – and find negative instructional effects but positive selection effects of students and/or teacher effort, yielding on average zero effects. In yet another review, Lloyd (1999) find positive effects of MA-classes on high ability students.

Using Swedish data, Sundell (2002) estimates the effect of MA-class-attendance in grade 2 on a number of abilities. Important to note is that the 752 students included in his study are not randomly sampled. When controlling for social and pedagogical background as well as initial achievements, the results show that students in MA-classes performed worse than other students in 12 out of 13 dimensions. The MA-students had for example lower mathematical ability, a less developed vocabulary and were perceived as more shy and troublesome by their teachers. However, they did perform better in reading comprehension.

3 Data

In this section, we describe the data used and show the differences between MA- and traditional classes in terms of some important aspects.

3.1 Data sources

Our main data source is a stratified panel data set: Student Panel 4, provided by Statistics Sweden.[5] In this panel, one cohort of students is followed through grade 3 to 9. In the first stage 35 municipalities were selected. In the second stage a random sample of grade 3 classes within these municipalities

[5] Participation in the study is voluntarily. About 4 percent of the originally sampled students were not able to or chose not to participate in the study.

were selected.[6,7] Within the selected classes, information from all students in grade 3 was collected. This means that for students in traditional classes, we have information on the whole class, while for students in MA-classes, we only have information on the part of the class that spends their third year in school in 1992. That is usually one half or one third of the class, depending on how the MA-class is constructed.

The sampling of grade 3 classes were done in 1992; hence most students are born 1982 and finish 9th grade in 1998. It is important to note that all students sampled in grade 3 are followed over time, regardless of whether they move or change class; hence, regarding these data there is virtually no attrition. The panel includes approximately 8500 individuals.

This panel data set is combined with additional register data from the data bases RAMS and LOUISE provided by Statistics Sweden. These data include socioeconomic background information such as parental education and immigrant status. Most of this information is measured in 1998. We focus on students who finish 9th grade the expected year 1998 or later.[8]

In addition, we have access to a survey with information on students, parents and teachers and their attitudes towards school-related issues. This information was collected when the students were in grade 6 by the Department of Education at Göteborg University.[9] Parents were asked about their involvement in school issues and if they actively had chosen school or simply accepted the nearest one. Teachers were asked about their work experience, whether they had a formal degree, and their attitude towards homework. Results from grade 6 cognitive tests of the students were also collected (a description of these tests is given in Appendix).

Due to non-response, survey information is only available for a sub-sample of the original sample. Of the individuals in the original sample, 85 percent have undertaken the grade 6 test, and 54 percent has answered all of the survey questions we use. It is this reduced sample we use for our analyses. Table A1 in Appendix shows the difference between the raw register data, data with test results available (the basic sample), and data with all survey information available (our survey sample). The differences in means are very small when comparing the raw data and the basic sample. In 6 out of 27 cases there are statistically significant differences at the ten percent level and in these cases the magnitudes of the differences are small. Comparing the raw data with the survey sample, there are some additional differences. The survey sample seems to consist of a slightly more

[6] Throughout the paper, we will show descriptive statistics and estimation results for unweighted data since the number of students in MA-classes is small and outliers could potentially be given large sampling weights.
[7] For more information about how the data was collected, see Statistics Sweden (1996)
[8] 16 students finished school one year earlier, but due to a changed grading system we do not include these in our sample.
[9] For a more detailed description of the data, see Härnqvist (2000).

"privileged" group of students than the raw data. For example, students in the survey sample have higher average credits and grade 6 test results, they are more seldom given special help or mother tongue education in grade 3, and their parents are better educated.

Our data are mainly collected at the individual level but some variables are for obvious reasons measured at the class level (teacher information) or school level (school information). The inference from descriptive statistics previously discussed (Tables in Appendix) is based on individual level variation. This potentially implies underestimated standard deviations for the class and school level variables. In Appendix we present descriptive statistics of these variables with standard deviations calculated on the school level (individual classes are unfortunately not observed in our data). At this level, there are no statistically significant differences between basic and survey sample.

3.2 Differences between MA-classes and traditional classes

Table 1 presents descriptive statistics for students in MA- and traditional classes in our survey sample. First of all, we can note that students in MA-classes have lower scores on the grade 6 cognitive tests. This could have two different explanations: one is that MA-classes are detrimental to student achievement, another is that we have negative selection into MA-classes. Regarding parental and student characteristics, the groups are relatively similar with two exceptions. Students in MA-classes have to a less extent mothers with university degree, and are more often given mother tongue education in grade 3.

Regarding teacher and class characteristics, the differences are more striking. MA-classes are usually smaller. The teachers in MA-classes are less experienced, have spent a shorter time in each class, and are more often on leave than teachers in traditional classes. The teachers' attitudes also differ[10]. Teachers in MA-classes put less emphasis on homework, basic knowledge and formal tests than teachers in traditional classes. MA-class-teachers also believe student influence to be more important than their colleagues in traditional classes. Hence, from these descriptive statistics it seems as if the pedagogical environment for students in MA-classes differs substantially from the environment in traditional classes.

[10] The attitude variables are measured on a 1-5 scale; the more important a teacher regards the issue, the higher the number. See Appendix for more details.

Table 1 Descriptive statistics of students attending an MA-class during grades 4–6 versus others, survey sample

	MA-class in grades 4–6		**Ordinary class in grades 4–6**	
	Mean	**Sd.**	**Mean**	**Sd.**
Individual characteristics				
Grade 9 credits	51.49	28.34	52.67	28.66
Grade 6 test results	46.99	28.23	52.20***	28.80
Female student	0.48	0.50	0.50	0.50
Early start	0.01	0.11	0.01	0.08
Late start	0.03	0.16	0.02	0.15
Birth month	6.11	3.43	6.27	3.35
Help in grade 3	0.16	0.37	0.19	0.39
Mother tongue in grade 3	0.11	0.31	0.08*	0.27
Non-Nordic student	0.07	0.25	0.06	0.24
Mother sec. educ.	0.46	0.50	0.46	0.50
Mother univ. educ.	0.26	0.44	0.33***	0.47
Father sec. educ.	0.37	0.48	0.40	0.49
Father univ. educ.	0.19	0.39	0.22	0.41
Father educ. miss	0.20	0.40	0.22	0.41
Mother educ. miss	0.06	0.23	0.06	0.23
Father non-Nordic	0.09	0.29	0.10	0.31
Mother non-Nordic	0.11	0.31	0.10	0.30
Birth country miss	0.00	0.06	0.00	0.03
Father birth country miss	0.03	0.18	0.03	0.16
Mother birth country miss	0.02	0.12	0.02	0.13
Parent attitude: active school choice	0.15	0.36	0.15	0.36
Parent attitude: parent help	1.87	0.89	1.91	0.95
Parent attitude: parent active	2.39	1.05	2.34	1.03
Teacher and class characteristics[1]				
International school	0.0000	0.0000	0.0007	0.0265
Confessional school	0.0032	0.0562	0.0028	0.0530
Special school	0.02	0.14	0.03	0.18
Grade 9 students	101.56	39.25	114.31***	40.42
Few grade 9 students	0.07	0.25	0.01***	0.11
Teacher experience	18.44	10.62	20.12***	9.65
Teacher not qualified	0.04	0.20	0.04	0.20
Class size	18.31	7.19	23.67***	5.91
Small class	0.13	0.33	0.01***	0.11
Large class	0.21	0.41	0.37***	0.48
Share boys	0.55	0.12	0.51***	0.11
Share Swe2 students	0.07	0.13	0.06	0.14
Teacher not full time	0.10	0.30	0.12	0.32
Teacher on leave	0.08	0.26	0.03***	0.17
Teacher year in class	2.56	1.28	2.80***	0.84
Teacher attitude: home works	3.56	0.94	3.83***	0.91

	Mean	Sd.	Mean	Sd.
Teacher attitude: tests	2.66	0.77	2.96***	0.94
Teacher attitude: basic knowledge	4.42	0.81	4.67***	0.60
Teacher attitude: student influence	3.96	0.80	3.85**	0.86
Teacher attitude: student responsibility	4.71	0.63	4.77*	0.50
N	317		4267	

Notes: Teacher and class information are collected at the individual level (the teacher has filled in one form for each student) and are treated as individual level information when calculating standard errors. The reason is that we cannot identify class in the data set. Significance levels for the difference in means: *** 1%, **5%, *10%.

4 Estimating the effect

In this section, we discuss issues of identification and describe our estimation strategy.

4.1 Identification

The potential effect on cognitive skills of attending an MA-class may stem from two types of factors: (i) the effect of interactions between students of different age and school experience and/or (ii) effects from parent and teacher involvement (see Figure 1). In the literature on MA-classes, most arguments for the beneficial effects of MA-classes focus on the student interaction effects.

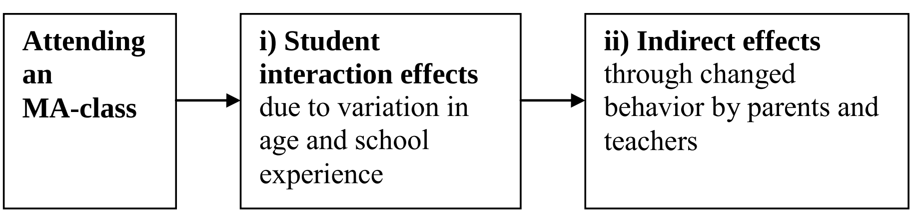

Figure 1 The different components of the MA-effects

Our purpose is to estimate the combined effect of (i) and (ii) on cognitive skills. This is the relevant question from a policy perspective and this is also what a randomized experiment would capture.[11] Our identification strategy is

[11] Although it is not the purpose of our paper, measuring the effect of interactions between students of different ages only (i.e. part of effect (i)) is relatively easily achieved. Given birth dates on every student within each class, we could simply estimate the effect of age variance within a class on student outcomes. In our data set we only have information on all students within each class for the traditional classes, and our sample size is much too small for a

a regression adjustment approach. This strategy relies on the selection on observables assumption, i.e., that the control variables we condition the analysis on, capture all potential selection of importance. In the following we discuss the credibility of this strategy for our context.

The type of selection is potentially affected by the reason for introducing MA-classes. There are basically two different reasons for why a student attends an MA-class: i) the number of students (or teachers) is too few to form an age homogeneous class and/or ii) a belief in the MA-concept.

i) mainly occurs in smaller municipalities far from larger cities and/or in schools with special profiles, such as confessional or international schools. This is problematic if municipality size/school type is correlated with important student/teacher characteristics (for example, if students in rural areas simultaneously have relatively lower abilities and are more likely to attend an MA-class.) Potential selection problems due to i) is mitigated by including municipality dummies, information on the size of the school (in grade 9) and information on whether the school has a special profile (e.g., international and confessional).

ii) implies selection problems if there is active sorting of students and/or teachers into MA-classes. This could happen for several reasons. Parents could actively choose class type for their child, either by direct class/school choice within their area of living or by moving to/from areas offering MA classes. Principals could place different types of students/teachers in MA-classes, or a different set of teachers could actively choose to work in MA-classes.

Regarding potential selection due to ii), we rely on our extremely detailed information, not only from register data but also from the survey data on students, parents and teachers and their self reported attitudes towards different school issues. One of our most important variables in this context is information on whether the parents have made an active school choice or not.

However, the choice of control variables is not straightforward, partly because the literature on MA-classes is rather vague, partly because some of our control variables are measured in grade 6. Although we view our variables to be controls for sorting and selection, it could be the case that some of them also reflect the indirect effect (ii) of attending an MA-class. One example is the variable attempting to measure how involved the parents are in school issues. Active parents may actively choose an MA-class (or traditional class) for their child (in which case the variable becomes an important control for selection), but it could also be the case that parents in

precise estimation of this effect. In spite of that, we find that the point estimate of the age variance in traditional classes on student achievement is negative.

MA-classes (or traditional classes) are forced to become more actively involved in school issues (in which case the variable represents the indirect effect of MA). We will estimate the effect of MA-classes both with and without the control variables considered potentially problematic. In the list of variables in Appendix, we have distinguished between these two types of variables.

4.2 Estimation strategy

We estimate the following model using ordinary least squares (OLS):

$$y = \alpha + \beta ma456 + \delta X_1 + \gamma X_2 + m + \varepsilon$$

where y denotes student achievement – either percentile ranked results from grade 6 cognitive tests or percentile ranked grade 9 credits. Our key explanatory variable, *ma456*, is a dummy variable for attending an MA-class all years in grades 4 to 6[12]. It is important to note that a class is defined as an MA-class only if it consists of students of both different ages and grades; this is not to be confused with traditional classes where some students happen to be born a different year than the others (for example, students with learning difficulties or especially skilled students). X_1 denotes the covariates used to control for selection bias. These include socioeconomic information such as parental education levels, immigrant status, gender and birth month of the student, and information on whether the student were given special help or mother tongue education in grade 3. For a complete list of all variables, see Appendix. When estimating the effect on grade 9 credits, we also control for the number of students in grade 9 at the school[13]. In addition, we have access to a variable indicating if the student attended an MA-class also during grades 7–9. This variable is included as a control in a separate estimation. X_2 denotes the variables used to control for selection, but where there is some uncertainty about whether or not they instead represent the indirect effects of MA-classes. These variables include the attitudes and behaviour of the teachers and parents. Finally, in all estimations we include municipality fixed effects, m.

We can also note that the two different measures of student outcomes, the grade 6 test results and the grade 9 credits, differ in two respects. Not only do they capture short- versus long run effects of attending an MA-class, they can also reflect slightly different types of skills. While the grade 9 credits are a weighted average of grades in different subjects, and as such could include

[12] Using other definitions of the explanatory variable ma456, such as a dummy for attending an MA-class only in grade 4 or at least one year during grades 4–6 or a cumulative variable capturing the number of years spent in an MA-class does not change the results.
[13] We do not have information on the size of the school in grade 6.

not only the teachers' assessment of the student's skills but also to some extent the students' behavior and diligence, the grade 6 tests are simply test results. The correlation between the two measures is also relatively low, with a correlation coefficient of 0.57.

Another thing to note is that we do not have information on whether the student attended an MA-class during grades 1–3. Since MA-class attendance in grades 1–3 is likely to be correlated with MA-class attendance in grades 4–6, it is possible that our dummy variable for MA-class attendance in grades 4–6 also partly captures the long run effects of earlier MA-class attendance.

With the estimation strategy above, we implicitly assume that the effect of attending an MA-class is equal for all groups of students. This may not be true – in fact, many of the arguments for or against MA-classes are concerned with how they affect different kinds of students. In particular, it is usually argued that attending an MA-class is especially valuable for students who do not perform as well as their peers. In many studies, it is shown that girls outperform boys in school and that immigrant students have lower school achievement than the average student. Hence, to relax the equal-effects assumption, we include interaction terms that allow the MA-effect to vary depending on gender, if the student has a non-Nordic background and if the student was low performing in grade 3 (measured by if the student were given special help in grade 3).

Finally, since some variables are common within classes and/or schools, inference is based on standard errors that allow for heteroscedasticity within schools. We cluster on school in grade 9 (as opposed to school in any earlier grade) due to data limitations – we only have information on the school in grade 9.[14]

5 Results

How does attending an MA-class affect student performance? In section 5.1 we estimate the average effect, while section 5.2 examine whether the effect varies by observed characteristics.

5.1 Main results

Table 2 and Table 3 show the effect of MA-classes on grade 6 cognitive tests and grade 9 credits, respectively. For the cognitive test results, there is a negative and statistically significant effect of attending an MA-class. The

[14] Since students may have attended different schools and/or classes during their school careers it is not obvious which school and/or class one would ideally want to use for clustering.

estimated effect on grade 9 credits is not statistically significant, although the point estimate is negative.

The magnitude of the effect is relatively large. Attending an MA-class in grades 4-6 reduces the cognitive test results by around 5 percentile points. This can be compared with the effect of class size reductions. In the Tennessee STAR experiment, reducing class size by one student increased student performance with almost one percentile point (Krueger, 1999).

The negative effect of attending an MA-class on grade 6 test results remains in about the same range regardless of the set of covariates used. A comparison between column 2 and column 3 shows no large differences. Hence, the variables added in column 3, that we view as good controls for selection but that potentially also could capture the indirect MA-effects, do not seem to be important in explaining the difference in achievement between students in MA- and traditional classes. This is interesting since these variables include the parental and teacher attitudes towards school issues. In Section 3 above, we noted that the largest differences between MA- and traditional classes were in terms of these different parental and teacher attitudes. At the same time, they seem unimportant for explaining the negative effect of MA-classes.

Table 2 OLS estimates of the effect of attending an MA-class during grades 4–6 on percentile ranked grade 6 test results, survey sample

	Grade 6 test results	**Grade 6 test results**	**Grade 6 test results**
MA grades 4–6	-5.681	-4.325	-4.509
	(2.059)***	(1.786)**	(1.825)**
Including X_1	No	Yes	Yes
Including X_2	No	No	Yes
R2	0.05	0.29	0.30
F-test if added parameters jointly equals zero		75.83	3.17
Probability>F		(0.0000)	(0,0026)
N	4584	4584	4584

Notes: All models include municipality dummies, standard errors in parentheses are clustered on schools. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

Table 3 OLS estimates of the effect of attending an MA-class during grades 4–6 on grade 9 credits, survey sample

	Grade 9 credits	**Grade 9 credits**	**Grade 9 credits**	**Grade 9 credits**
MA grades 4–6	-2.579	-0.989	-1.169	-0.915
	(1.742)	(1.317)	(1.315)	(1.336)
Including X^1	No	Yes	Yes	Yes
Including X^2	No	No	Yes	Yes
Including MA grades 7–9	No	No	No	Yes
R2	0.04	0.30	0.30	0.31
F-test if added parameters jointly equals zero		53.76	4.35	23.11
Probability>F		0.0000	0.0019	0.0000
N	4584	4584	4584	4584

Notes: All models include municipality dummies. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

5.2 Heterogeneous effects

Table 4 shows the results from the heterogeneous effects estimations. All interaction terms are very imprecisely estimated and we find no statistically significant differences for any of the subgroups studied – girls, immigrants or low-performing students. This means that we cannot find support for the arguments commonly used in favour of MA-classes – that MA-classes especially should benefit low performing students.[15]

[15] We have also studied the same heterogeneous effects on the grade 9 credits but find no statistically significant differences between groups.

Table 4 OLS estimates of heterogeneous effects of attending an MA-class during grades 4–6 on percentile ranked grade 6 test results, survey sample

	Grade 6 test results
MA grades 4–6	-6.917
	(2.907)**
Female student	-3.278
	(0.809)***
(MA grades 4–6)* (Female student)	4.730
	(3.959)
Help grade 3	-24.891
	(0.916)***
(MA grades 4–6)* (Help grade 3)	-1.555
	(3.842)
Non-Nordic student	-6.722
	(2.510)***
(MA grades 4–6)*(Non-Nordic student)	7.151
	(5.709)
Including X_1	Yes
Including X_2	Yes
R2	0.30
N	4584

Notes: All models include municipality dummies. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

6 Sensitivity analysis

Our regression adjustment approach is based on several assumptions such as correct functional form, no omitted variables and no selection into class type conditional on the observable variables. In general, these assumptions are strong, although the richness of control variables increases the credibility of this study. Below, we extend the main analysis in different ways in order to investigate the robustness of our results. First, we include an additional proxy variable for ability and re-estimate the main model. Second, we re-estimate the model using a larger sample and show calculations on the probability of entering the sample, depending on ability and class type. Finally, we also employ a propensity score matching method.[16] All in all, these sensitivity analyses do not modify our conclusion that attending an MA-class is negative for student achievement.

An additional issue is potential measurement errors. Since our most important control variables stem from register information, we do not consider this to be a serious problem. MA-attendance is a survey variable but

[16]Matching approaches relaxes the linearity assumption. In addition, matching addresses the issue of selection on observables somewhat differently by only comparing individuals within the common support. See for example Black and Smith (2004) for a discussion of this in the educational context.

it is hardly a sensitive question. Hence, we do not suspect measurement errors. In addition, if there are, it would only imply attenuation bias, yielding a smaller-than-true coefficient estimate.

6.1 Adding a proxy for ability

If there is selection of students and/or teachers into different class types depending on unobserved ability, the estimated effect of attending an MA-class is biased. To investigate this, we add the grade 6 test result for spatial ability ("metal folding")[17,18] to the fully extended model (Table 2, column 3). Recent evidence in Öckert (2009) suggests that spatial ability is less malleable to schooling than inductive or verbal ability, which is our reason for not using it as an outcome variable. However, it might indeed serve as a proxy for initial ability.

Including spatial ability in the model reduces the MA-coefficient slightly (from -4.509 to -3.169) but the difference is not statistically significant. Hence, unobserved heterogeneity in terms of ability does not seem to bias our results.

In addition, and in order to explicitly allow the MA-effect to vary depending on spatial ability, we also include an interaction term between this ability measure and the MA-coefficient. The estimate of this interaction term is not statistically significant and the parameter of interest (the MA-coefficient) is about the same as in the main estimations (in fact slightly larger but less precisely estimated).

6.2 Sample selection analysis

Sample selection is a concern for validity, both external and internal[19]. If the ability distribution differs between students who respond to the survey and the register sample, external validity could be violated. If the ability distribution of survey respondents differs between MA-classes and traditional classes, internal validity could be threatened.[20] Some calculations can help shed light on this issue.

Unconditional on covariates, there is a small positive correlation between ability (as measured by the variable "metal folding", our proxy for initial ability) and the propensity to answer the survey (the response rate). This

---

[17] See Appendix for more information about this test.

[18] Ideally, we would want our measure of ability to be collected prior to school start or at least prior to our period of observation but such a variable is not available.

[19] In this context, external validity focuses on the relationship between the survey sample and the register sample. Since we use an unweighted stratified data set, the results are not neccessarily externally valid in the sense that they reflect the effect for the total Swedish population.

[20] Among MA-students 55 percent answered the survey while the corresponding number for comparison students is 64 percent.

means that our sample consists of a slightly more high-ability group of students than the register data, which may violate external validity. Moreover, the magnitude of this correlation also differs between students in different class types – the correlation is 0.036 for MA-students and 0.049 for students attending traditional classes. In practice, this means that we draw comparison students slightly more from the upper part of the ability distribution which may violate internal validity.

One way to examine this issue is to re-estimate the model on a larger sample closer to the register sample. We use the basic sample, which only requires register covariates to be available, and hence is a more representative sample than the survey sample[21]. The results are shown in Table A2 in Appendix. Clearly, the estimated effect of attending an MA-class is still negative, although slightly less so, in the basic sample compared to the survey sample.

In addition, we estimate how the probability of being included in the survey sample depends on MA-class attendance, ability and the interaction between MA-class and ability, plus the register covariates[22]. The results are shown in Table 5 below. Clearly, the propensity to answer the survey is slightly lower among MA-students; however, this in itself does not necessarily violate external or internal validity. The coefficient for spatial ability is close to zero (although statistically significant) which indicates that conditional on covariates, our survey sample is representative of the total population in terms of ability. The interaction term between MA-status and ability is also zero and statistically insignificant, indicating that conditional on covariates, we have no selection in terms of ability into different types of classes. In sum, the register control variables seem to handle the potential problems of sample selection well.

[21] The basic sample contains 7234 observations, compared to the 8531 individuals in the raw data.

[22] The model becomes $P(\text{included in the survey sample})_i = a + bMA_i + cABILITY_i + dMA_i *ABILITY_i + e*X_i + f_i$, where X contains the register covariates. The coefficients of interest are c (to investigate sample selection threatening external validity) and d (to investigate sample selection threatening internal validity).

Table 5 Sample selection into survey sample, on basic sample

	Grade 6 test results
MA grades 4-6	-0.092
	(0.054)*
Spatial Ability	0.000
	(0.000)*
Spatial Ability*MA	-0.000
	(0.001)
R2	0.07
N	7228

Notes: The basic sample is reduced by 6 observations due to missing information on the ability measure. The model includes register covariates and municipality dummies. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

6.3 Propensity score matching

Finally, we also estimate the effect of MA-class attendance by propensity score matching. First, we estimate a probit regression model for the probability of attending an MA-class. In the probit regression, we use the fully extended model (column 3, Table 2). Second, we match (nearest neighbor without replacement) on these predicted values so that we for each MA-student get one comparable individual who has attended a traditional class. In this matched sample, there are no statistically significant differences between MA- and comparison students. Using this more homogeneous sample, we estimate the effect of MA without any covariates, i.e., we simply compare the mean values. The results are presented in Table 6. The point estimate of attending an MA-class is about the same as in the linear regression model but, as expected due to a smaller sample size, the precision is lower. Hence, using a matching model which relaxes the linear functional form of OLS and also excludes non-comparable individuals (outside the common support) does not alter our main conclusions.

Table 6 Matching approach, survey sample

	Grade 6 test results	Grade 9 credits
MA grades 4–6	-3.532	-0.479
	(3.412)	(2.658)
R2	0.00	0.00
N	608	608

Notes: The sample is slightly reduced (by 2*15 individuals) since the probit regression perfectly predicts success/failure for 15 MA-students. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

7 Concluding remarks

Despite ambiguous scientific evidence, mixed-age classes are a common phenomenon in schools around the world. In some cases, it is because of demographic necessity; in other cases, it is because MA-classes are claimed to enhance student achievement. In Sweden these types of classes have been rapidly re-introduced and nowadays, around one fourth of all children attend an MA-class during grades 4–6.

In this paper, we present evidence that MA-classes have a negative effect on short-term cognitive skills, as measured by the grade 6 cognitive tests. This effect is robust to a rigorous sensitivity analysis and we cannot detect statistically significant differences in this effect for girls, low performing students or students with a non-Nordic background. The effect of attending an MA-class on grade 9 average credits is not statistically significant, although the point estimate is negative.

We have not been able to distinguish between MA-classes introduced out of pedagogical beliefs and MA-classes introduced out of economic and/or demographic necessity. Since the effect of MA-class attendance could differ between these two groups, this would be an interesting topic for future research.

References

Andrae Thelin, A. (1991): Undervisning i åldersblandade grupper. En granskning i samarbete mellan skolöverstyrelsen och länsskolnämnderna läsåren 1989/90 och 1990/91. Skolöverstyrelsens rapport R91:32

Black, D. A. and J. Smith (2004): How robust is the evidence on the effect of college quality? Evidence from matching, Journal of Econometrics, Elsevier, vol. 121, no. 1-2

Edlund, A. C and K. Sundell (1999): Åldersintegrerat eller åldersindelat? En jämförande studie av 1 111 elever i årskurs, FoU-rapport 1999:9, Stockholm socialtjänstförvaltning.

Hanushek, E. (1999): Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects, Educational Evaluation and Policy Analysis, vol. 21, no. 2

Härnqvist, K. (2000): Evaluation through follow up: a longitudinal program for studying education and career development, in Jansson (ed.): Seven Swedish longitudinal studies in the behavioral sciences, report no 2000:8, Swedish council for planning and coordination of research.

Krueger, A. (1999): Experimental estimates of education production functions, Quarterly Journal of Economics, vol. 114, no. 2

Krueger, A. (2003): Economic considerations and class size, Economic Journal, vol. 113, no. 485

Little, A.W. (2001): Multigrade teaching: towards an international research and policy agenda, International Journal of Educational Development, vol. 21, no. 6

Lloyd, L. (1999): Multi-age classes and high ability students, Review of Educational Research, vol. 69, no. 2.

Mason, A. & R.B. Burns (1996): 'Simply no worse and simply no better' may simply be wrong: a critique of Veenman's conclusion about multigrade classes, Review of Educational research, vol. 66, no. 3
Mason, A. & R.B. Burns (1997): Reassessing the effects of combination classes, Educational Research and Evaluation, vol. 3, no. 1
Sandqvist, K. (1994): Åldersintegrerad undervisning: En kunskapsöversikt, HLS, Stockholm.
Statistics Sweden (1996): Elevpanel för longitudinella studier, Elevpanel 4, Statistiska meddelanden U73, SM 9601, Örebro.
Sundell, K. (1995): Åldersindelat eller åldersblandat? Forskning om ålderssammansättningens betydelse i förskola och grundskola, Studentlitteratur, Lund.
Sundell, K. (2002): Är åldersblandade klasser bra för eleverna? En jämförande studie av 752 elever i årskurs 2 och 5, FoU-rapport 2002:7, Stockholm socialtjänstförvaltning.
Svensson, A. (1964): Sociala och regionala faktorers samband med över- och underprestation i skolarbetet: Pedagogisk-sociologiska studier jämte en beskrivning av skolstatistikens individualuppgifter, Department of Education, Göteborg university.
Veenman, S. (1995): Cognitive and non cognitive effects of multigrade and multi-age classes: a best-evidence synthesis, Review of Educational Research, vol. 65, no. 4
Vinterek, M. (2001): Åldersblandning i skolan - elevers erfarenheter, doctoral thesis, Umeå University.
Vinterek, M. (2003): Åldersblandade klasser. Lärares föreställningar och elevers erfarenheter, Studentlitteratur, Lund.
Öckert, B. (2009): Does going to high school make you smarter?, Manuscript, Institute for Labour Market Policy Evaluation (IFAU)

Appendix

A1 List of variables

List of variables: Register data

Variable name	Definition
Grade 6 test results	Percentile rank of the sum of the scores on the tests in number series and opposites given in grade 6
Grade 9 credits	Percentile rank of a summary measure of the student's 16 best credits in grade 9
MA grades 7–9	A dummy that equals 1 if the students attends an MA-class in any grade between grades 7–9
MA grades 4–6	A dummy that equals 1 if the students attends an MA-class during grades 4-6
Municipality dummies	One dummy for each municipality
Female student	A dummy that equals 1 if the student is female
Early start	A dummy that equals 1 if the student is born after 1982
Late start	A dummy that equals 1 if the student is born before 1982
Birth month	The student's month of birth
Help in grade 3	A dummy that equals 1 if the student has been given any form of special education intended for low performing students ("särundervisning", "anpassad studiegång" or "specialundervisning på annat sätt") in grade 3
Mother tongue in grade 3	A dummy that equals 1 if the student attended mother tongue education in grade 3
International school	A dummy that equals 1 if the school has an international profile
Confessional school	A dummy that equals 1 if the school has a confessional profile
Special school	A dummy that equals 1 if the school is not ordinary, for example schools at hospitals
Grade 9 students	The number of students in grade 9 at the school, collected in grade 9
Few grade 9 students	A dummy that equals 1 if the number of students in grade 9 at the school is smaller than 30
Non–Nordic	A dummy that equals 1 if the student is born in a non-Nordic country (missing values are set to 1)
Mother secondary education	A dummy that equals 1 if the mother of the student has secondary education, at most 5 years in addition to compulsory schooling. (Missing values are set to 0).
Mother university education	A dummy that equals 1 if the mother of the student has university education, more than 5 years in addition to compulsory schooling. (Missing values are set to 0).
Father secondary education	A dummy that equals 1 if the father of the student has secondary education, at most 5 years in addition to compulsory schooling. (Missing values are set to 0).
Father university education	A dummy that equals 1 if the father of the student has university education, more than 5 years in addition to compulsory schooling. (Missing values are set to 0).
Mother non-Nordic	A dummy that equals 1 if the mother of the student is born in a non Nordic country (missing values are set to 1)

Father non-Nordic	A dummy that equals 1 if the father of the student is born in a non Nordic country (missing values are set to 1)
Birth country missing	A dummy that equals 1 if information about the student's country of birth is missing
Mother education missing	A dummy that equals 1 if information about the mother's education is missing
Father education missing	A dummy that equals 1 if information about the father's education is missing
Father birth country missing	A dummy that equals 1 if information about the father's country of birth is missing
Mother birth country missing	A dummy that equals 1 if information about the mother's country of birth is missing

List of variables: Survey data collected in grade 6

Variable name	Question	Definition
	To teachers:	
Teacher experience	What is your teacher experience in years?	A variable ranging from 1 to 43 (measured in years)
Teacher not qualified	Do you have a certificate qualifying you to teach at this level?	A dummy that equals 1 if the answer is no
Class size	What is the number of girls and boys in the class?	The sum of boys and girls in the class - ranging from 0 to 60
Small class	Constructed from Class size	A dummy that equals 1 if the class size is smaller than 10
Large class	Constructed from Class size	A dummy that equals 1 if the class size is larger than 25
Share boys	Constructed from Class size	The share of boys in the class
Share Swe2 students	What is the number of students in your class that take Swe2? Constructed from Class size	The share of students in the class taking a special course in Swedish adapted for students who do not have Swedish as mother tongue
Teacher not full time	Do you work full time?	A dummy that equals 1 if the teacher work part time, 0 if full time
Teacher on leave	Have you been on leave during the last year?	A dummy that equals 1 if the teacher has been on leave full time or part time
Teacher year in class	How many years have you taught this class?	A variable ranging from 1 to 8 (measured in years)
Teacher attitude: home works (belongs to X_2, the extended set of covariates)	How important are home works and oral tests?	A variable ranging from 1 to 5 in the following way: Very important 5, Rather important 4, In between 3, Rather unimportant 2, Not at all important 1

Teacher attitude: tests (belongs to X_2, the extended set of covariates)	How important are formal tests?	See above.
Teacher attitude: basic knowledge (belongs to X_2, the extended set of covariates)	How important is the emphasis of basic skills?	See above.
Teacher attitude: student influence (belongs to X_2, the extended set of covariates)	How important is student influence during planning?	See above.
Teacher attitude: student responsibility (belongs to X_2, the extended set of covariates)	How important is it that the student takes own responsibility?	See above.
	To parents:	
Parent attitude: active school choice (belongs to X_2, the extended set of covariates)	Have you chosen another than the closest school to your child?	A dummy that equals 1 if the answer is *yes* or *yes we are going to* and 0 if *no or doubtful (probably not)*
Parent attitude: parent help (belongs to X_2, the extended set of covariates)	Do you participate in your child's school work?	A variable ranging from 1 to 5 in the following way: Very often 5 Rather often 4 Sometimes 3 Rarely 2 Almost never 1
Parent attitude: parent active (belongs to X_2, the extended set of covariates)	How much do you participate in school activities?	See above.

A2 The cognitive tests from grade 6

There are three test scores from grade 6 available. The tests represent verbal, spatial and reasoning abilities and are called: Opposites (motsatser), Number series (talserier) and Metal folding (platvik). In the test called Opposites, the child is asked to find the opposite of a given word among four choices (40 items, 10 minutes). In the Number series test the child is instead asked to complete number series (40 items, 18 minutes). In the last test, Metal folding, the child is asked to find the three-dimensional object among four choices that can be made from a flat piece of metal (40 items, 15 minutes).[23] The results on each of these tests are measured on a scale ranging from 0 to 40.

In this paper, we use the percentile ranked sum of two of the tests, Opposites and Number series. The correlation coefficient between each of these tests and the grade 9 credits is 0.51. The third test, Metal folding, involves tasks not regularly practiced in schools, and its correlation coefficient to the grade 9 credits is 0.36.

A3 Additional tables

Table A1a Register data, Basic sample and Survey sample, variation at individual level

	Register data			Basic sample, N=7,234		Survey sample, N=4,584	
	Mean	**Sd.**	**N**	**Mean**	**Sd.**	**Mean**	**Sd.**
Key variables							
Grade 9 credits 1)	202.17	59.88	8,490	203.93*	58.56	209.16***	56.77
Grade 6 test results 2)	44.24	12.68	7,420	44.32	12.66	45.13***	12.55
MA grades 4–6	0.08	0.26	8,531	0.08	0.27	0.07	0.25
MA grades 7–9	0.03	0.17	8,531	0.02**	0.15	0.02***	0.15
Covariates capturing selection							
Female student	0.49	0.50	8,515	0.49	0.50	0.50	0.50
Early start	0.01	0.09	8,515	0.01	0.09	0.01	0.09
Late start	0.03	0.16	8,515	0.02	0.15	0.02	0.15
Birth month	6.28	3.36	8,515	6.27	3.36	6.26	3.35
Help in grade 3	0.21	0.40	8,531	0.20	0.40	0.19***	0.39
Mother tongue in grade 3	0.10	0.30	8,531	0.09	0.29	0.08***	0.27
International school	0.0011	0.03	8,360	0.0007	0.0263	0.0007	0.0256
Confessional school	0.0054	0.07	8,360	0.0043	0.0653	0.0028**	0.0532
Special school	0.04	0.19	8,360	0.04	0.19	0.03*	0.18
Grade 9 students	113.80	41.58	8,331	113.52	41.09	113.42	40.47
Few grade 9 students	0.03	0.16	8,331	0.02*	0.14	0.02***	0.13
Non-Nordic student	0.07	0.26	8,531	0.07	0.25	0.06**	0.24
Mother sec. educ.	0.45	0.50	8,531	0.45	0.50	0.46	0.50
Mother univ. educ.	0.30	0.46	8,531	0.30	0.46	0.32***	0.47
Father sec. educ.	0.38	0.48	8,531	0.39	0.49	0.40***	0.49
Father univ. educ.	0.20	0.40	8,531	0.20	0.40	0.22**	0.41
Father educ. miss	0.26	0.44	8,531	0.24*	0.43	0.22***	0.41
Mother educ. miss	0.07	0.26	8,531	0.07	0.25	0.06***	0.23
Father non-Nordic	0.12	0.33	8,531	0.11*	0.32	0.10***	0.30

[23] A more detailed description of the test scores are given by Svensson (1964).

	Register data			Basic sample, N=7,234		Survey sample, N=4,584	
	Mean	Sd.	N	Mean	Sd.	Mean	Sd.
Mother non-Nordic	0.12	0.32	8,531	0.11*	0.31	0.10***	0.30
Birth country miss	0.0014	0.04	8,531	0.0011	0.0332	0.0011	0.0330
Father birth country miss	0.03	0.18	8,531	0.03	0.17	0.03	0.16
Mother birth country miss	0.02	0.13	8,531	0.02	0.13	0.02	0.13
Teacher experience						20.00	9.73
Teacher not qualified						0.04	0.20
Class size						23.30	6.16
Small class						0.02	0.14
Large class						0.35	0.48
Share boys						0.51	0.11
Share Swe2 students						0.06	0.14
Teacher not full time						0.12	0.32
Teacher on leave						0.03	0.18
Teacher year in class						2.78	0.88
Parent attitude: active school choice						0.15	0.36
Covariates capturing selection and/or indirect MA-effects							
Teacher attitude: home works						3.82	0.91
Teacher attitude: test						2.94	0.93
Teacher attitude: basic knowledge						4.65	0.62
Teacher attitude: student influence						3.86	0.85
Teacher attitude: student responsibility						4.76	0.51
Parent attitude: parent help						1.90	0.95
Parent attitude: parent active						2.34	1.04

Notes: 1) and 2) are not percentile ranked. Significance levels for comparisons against register data: *** 1%, **5%, *10%.

Table A1b Basic sample and Survey sample, variation at the school level

	Basic sample, N=484 (schools)		Survey sample, N=375 (schools)	
	Mean	**Sd.**	**Mean**	**Sd.**
International school	0.01	0.09	0.01	0.07
Confessional school	0.02	0.14	0.01	0.11
Special school	0.05	0.21	0.04	0.20
Grade 9 students	98.27	45.81	102.50	44.43
Few grade 9 students	0.08	0.28	0.06	0.24
Small class			0.02	0.13
Large class			0.34	0.41
Teacher experience			19.51	8.46
Teacher not qualified			0.04	0.15
Class size			23.29	5.89
Share boys			0.50	0.09
Share Swe2 students			0.09	0.15
Teacher not full time			0.12	0.26
Teacher on leave			0.05	0.18
Teacher year in class			2.72	0.71
Teacher attitude: home works			3.89	0.76
Teacher attitude: test			2.98	0.78
Teacher attitude: basic knowledge			4.62	0.54
Teacher attitude: student influence			3.89	0.72
Teacher attitude: student responsibility			4.77	0.44

Table A2 OLS-estimates of the effect of attending an MA-class during grades 4-6 on percentile ranked grade 6 test results, controlling for register covariates only

	Basic sample	Survey sample
	Grade 6 test results	**Grade 6 test results**
MA grades 4–6	-1.171	-2.141
	(0.558)**	(0.795)***
Including register covariates	Yes	Yes
Constant	44.518	44.672
	(0.533)***	(0.643)***
R2	0.31	0.30
N	7234	4584

Notes: All models include municipality dummies. Significance levels: *** 1%, **5%, *10%. Standard errors in parentheses, clustered on schools.

Essay 2: Do pre-school interventions further the integration of immigrants? Evidence from Sweden[1]

(together with Peter Fredriksson, Caroline Hall and Per Johansson)

1 Introduction

Immigrant students typically perform substantially worse than native students in the OECD countries. According to PISA, the performance gap between first generation immigrants and natives amounts to around half a standard deviation in math, reading, and science (OECD 2006a). The achievement gaps between immigrants and natives are particularly large in Middle and Northern Europe (Schneeweis, 2009).

The size of the achievement gaps across countries depends on the characteristics of immigrants; in particular, immigrant source countries are likely to be important. But the characteristics of (host-country) educational institutions should also matter. It is intuitively plausible that pre-primary education is one important factor. Indeed, Schneeweis (2009), in her analysis of aggregate cross-country data, found that immigrant/native achievement gaps are lower in countries that make extensive use of pre-primary education.

The main contribution of this paper is that we directly examine whether pre-primary interventions reduce the immigrant/native gap in school performance. We use individual data containing information on childcare attendance, measures of cognitive achievement at age 13, and long-run educational attainment.[2] We thus examine the medium and long-run effects of pre-primary interventions.

[1] We thank Tuomas Pekkarinen for very helpful suggestions. We also thank seminar participants at the IFAU and Uppsala university for useful comments.

[2] The data come from the so-called UGU-project which is run by the Department of Education at Göteborg University; see Härnqvist (2000) for a description of the data. To these data we have matched educational attainment from the Educational Register (*Utbildningsregistret*) maintained by Statistics Sweden.

We are thus contributing to the recent flurry of papers analyzing the effects of (universal) pre-school interventions; see, e.g., Baker et al. (2008), Berlinski et al. (2009), Datta Gupta and Simonsen (2007), Gormley and Gayer (2005), and Havnes and Mogstad (2009). The literature has examined both cognitive and non-cognitive outcomes. The findings are mixed. Studies focusing on cognitive outcomes tend to find positive short-run effects, but the analysis in Magnuson et al. (2009) suggests that these may dissipate in the medium run. Studies focusing on short-run non-cognitive outcomes suggest that the effects may be negative, at least as indicated by parents (Baker et al., 2008); little is known about the longer-run effects on non-cognitive outcomes.[3] Thus, most studies of universal pre-school have focused on short-run effects. Havnes and Mogstad (2009), however, is a recent exception.[4] They find substantial positive effects of pre-school attendance on long-run education attainment. Apart from Schneeweis (2009) we have seen no other paper focusing on immigrants.

Pre-school interventions are likely to reduce inequality in education performance if the alternative to pre-schools (usually the home-environment) is worse for disadvantaged children than for advantaged children. For comparison, we also provide estimates for children with low-educated parents. We thus examine whether any effects are particular to immigrants or whether they apply to disadvantaged groups in general.

Our data cover cohorts born between 1967 and 1982. The time period spanned by these data involve changes in policy which have affected female labor supply and the demand for childcare. The past 40 years have seen a remarkable rise of female labor force participation in Sweden which is intimately tied to an increase in childcare enrolment.[5] The increase in female participation rates and the build-up of pre-schools/childcare were partly the responses to a tax reform in 1971. In 1971, the tax system changed from family taxation to individual taxation. This reform improved the incentives for women – particularly high-skilled women – to enter the labor market.

We are interested in the question of how childcare attendance affects the cognitive achievement gap between immigrants and natives in the medium and the long run. Ideally we would have liked to estimate mean impact of

[3] Note that this statement pertains to the effects of universal childcare/pre-schools. The studies of the Perry Preschool and Abecedarian programs suggest substantial and favorable longer run effect on non-cognitive (behavioral) outcomes for the particularly disadvantaged groups that participated in these experiments; see Karoly et al. (2005).

[4] See Jonsson (2004) for a study on the effects of pre-schools on educational attainment using Swedish data.

[5] Daycare centers/pre-schools have both caring and school preparatory elements. The official terminology changed from daycare to pre-schools in 1998 when a curriculum was introduced. Note that children in daycare/pre-schools have always been "taught" by staff with some pedagogical training. In the sequel we try to adhere to the following terminology. We use "childcare" to refer to both "pre-schools" and "family daycare"; the latter two concepts are defined more closely in the next section.

pre-school attendance as well. But the fact that pre-school attendance is so intimately tied to female labor force participation makes such an analysis much harder. We will rely on a selection on observables assumption to estimate the achievement gaps between immigrants and natives. We perform sensitivity analyses to evaluate the credibility of this assumption. Our conclusion from the sensitivity analyses is that the impact of childcare attendance on the achievement gaps seems credibly identified.

We find that childcare attendance reduces the gap in language skills between children from immigrant backgrounds relative to native children. We find no differential effects by mother's education, however. Nor does pre-school education affect the distribution of inductive skills or long-run educational attainment.

2 Background facts

The purpose of this section is to provide some background facts. We provide these facts along three dimensions: first, we describe childcare, its expansion and the nature of the "treatment"; second, we describe the evolution of female labor supply; and, third, we describe how the composition of children enrolled in childcare has evolved over time.

2.1 Childcare – expansion, content and alternatives

Prior to the late 1960s, childcare was available on a small scale and distinctively targeted at disadvantaged children. The words of a public committee (SOU 1944:20) illustrate the prevailing view. The committee advocated the introduction of pre-schools arguing that "Children from disadvantaged backgrounds should have the possibility to spend time in an activity that furthers their development. [Therefore] pre-schools should be introduced, where children through play (and other activities) enhance social skills, perception, and verbal skills". This policy prescription has been echoed by Heckman (and coauthors) in a series of papers (e.g., Cunha et al. 2006).

A major change in tax policy in 1971 changed the composition of children enrolled in childcare substantially. The policy reform moved income taxation from joint to individual taxation. The tax reform improved the incentives for women (typically the second earners) to enter the labor market, since marginal income tax rates were reduced substantially. In fact, the reform was preceded by the introduction of optional individual taxation in 1966, where couples could move to individual taxation if this minimized total tax payments (Selin, 2008). This policy change seems to have spurred the demand for childcare. Pre-school enrolment rates started to increase in the second half of the 1960s; see Figure 1.

Since the late 1960s there has been an impressive increase in pre-school enrolment. In 1970, 4.5 percent of children aged 1–5 were enrolled in pre-schools. By 1985, the share had increased to 32 percent and by 2007 it had increased further to 80 percent.

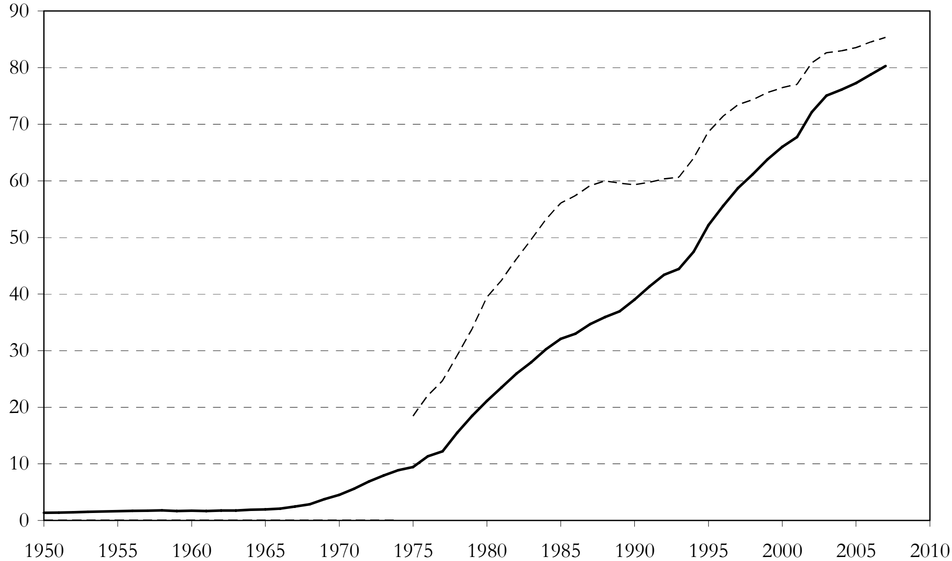

Figure 1 Share of population aged 1–5 enrolled in pre-schools (solid) and pre-schools plus family day-care units (dashed), percent, 1950-2007

Notes: From 1975-2007, pre-school enrolment is reported by age. Before 1975 only total enrolment is available. We have used 1975 data on the share of children above age 5 and below age 1 to adjust the pre-1975 data. Pre-1968, there is only information on the number of slots in pre-schools. We have used the relationship between the number of slots and the number of enrolled children in 1968 to adjust the pre-1968 data.
Sources: Statistics Sweden (*Utbildningsstatistisk Årsbok*, 1978, 1999, 2002, 2009; *Befolkningsförändringar*, 1950-1967, *Befolkningen* 1968-2007).

In terms of the increase in the total number of children involved in childcare activities the solid line is somewhat misleading. Since 1970, so called family daycare units have been available. By the mid 1980s, these daycare units hosted a substantial share of children in the pre-school ages. In 1985, 56 percent were enrolled in some childcare activity (either pre-schools or family daycare units); see the dashed line in Figure 1.

The municipalities provide for both pre-schools and family daycare. Pre-schools are organized facilities, with regular opening hours, while family daycare takes place in private homes. In order to shed light on the nature of treatmen, we provide some information on, *inter alia*, resources, staffing and staff qualifications at pre-schools and family daycare.

Relative to the rest of OECD (see OECD 2009a), expenditures on pre-primary education appears to be about average. For example, expenditure per student relative to GDP per capita was slightly below average in 2006 while expenditure per student in PPP converted US Dollars was slightly above average. Looking instead at the number of children per teacher (the student/teacher ratio) this was below the OECD average in 2006: 12.5 students per teacher in Sweden while the OECD average amounted to 14.9.

How has the student/teacher ratio evolved over time? The available data (see Johansson and Åstedt, 1996) suggest no major changes over time. The number of students per staff was almost the same in 1970 as in 1994. The most relevant period for our purposes, however, is the period 1970–85. During this period there seems to have been a slight reduction in the student/teacher ratio.

Basically, there are two kinds of employees in Swedish pre-schools: teachers and child minders. Pre-school teachers have tertiary education (currently 3.5 years) while child minders, at the time, had 2 years of upper-secondary education. In 1980, 45 percent of all employees had pre-school teacher training while 46 percent had child minder training (see Johansson and Åstedt 1996). Between 1970 and 1990, there appear to have been no major changes in the relative shares of pre-school teachers and child minders.

The fact that almost half of the staff employed by pre-school have pedagogical training arguably suggests that pre-school activities have (and have had) pedagogical content. In fact, the first Kindergarten was established in the late 1800s.[6] Another hallmark was the public commission (*Barnstugeutredningen*) established in 1968. The public commission had a distinct developmental psychologist or educationalist approach. The work of the commission eventually led to the law on public pre-schools (*Lag om allmän förskola*) which was implemented in 1975. At that time, pre-school activities were the responsibility of the Ministry of Health and Social Affairs and monitoring was conducted by the National Board of Health and Welfare, which, *inter alia*, issued pedagogical guidelines. In 1996, the responsibility for pre-schools was transferred to the Ministry of Education and in 1998 a national pre-school curriculum was introduced. While the transfer of responsibility to the Ministry of Education may have been an important signal, there is little evidence to suggest that the pedagogical emphasis changed. Relative to the earlier guidelines and pre-school programs issued by the National Board of Health and Welfare, the curriculum emphasized (in general terms) the pedagogical goals, rather than how they should be attained.

In a couple of reports the OECD has compared early childhood education and care across a selection of countries; see OECD (2001, 2006b). The OECD (2001) emphasizes that Swedish pre-schools appear to be of high-

[6] Richardson (2004) describes the historical evolution of the Swedish schooling system. The remainder of the text draws on this source unless explicitly stated.

quality: the fraction of pre-school teachers with tertiary education is high, almost the entire staff is trained to work with children, and child/staff ratios are low. We have not been able to detect any major changes in these quality indicators since 1970. Therefore, we infer that the relative quality of Swedish pre-schools is likely to have been high during the 1970s and 1980s – the time period most relevant to our empirical analysis – as well.

Family daycare units have in common with regular pre-schools that they offer an environment which is different from the home environment. Thus, for example, it is more likely that immigrant children interact with native born individuals in both pre-schools and family daycare than in the home environment.

But family daycare units differ somewhat from pre-schools in other respects. Family daycare units are typically staffed by individuals without (tertiary) pedagogical training; nevertheless, the providers are typically trained to take care of children.[7] Furthermore, by construction, group sizes are smaller in the daycare units than in regular pre-schools.

2.2 The evolution of female labor supply

The tax reform (alluded to above) improved the incentives for women to participate on the labor market. The 1970s saw other changes which may have contributed to increasing female labor supply. In 1974, a parental leave system was introduced. The system involved parental leave compensation which was proportional to individual earnings (prior to child birth) up to a ceiling. The policy created an incentive to enter the labor market prior to giving birth. During the 1970s, there was also a rather dramatic reduction in wage dispersion. Since men were usually the prime wage-earners, wage compression may have induced an increase in labor supply among married women.

All in all, the changes during the 1970s improved the incentives for women to supply labor. Figure 2 illustrates how married women responded to the change in tax policy. In the mid 1960s, 50 percent of married women aged 25-54 participated in the labor force. Following tax policy changes, changes in parental leave policy (foremost the introduction of a parental leave system in 1974) and the build-up of childcare, the participation rate among married women converged to the participation rate among single women by the second half of the 1980s.[8]

Female participation rates in Sweden are among the highest in the world. In 2008, 87.5 percent of the female population aged 25–54 participated in

[7] OECD (2001) reports that 72 % of family daycare providers have either a child minder certificate or have taken a mandatory child minder course from their municipal employers.

[8] We cannot update Figure 2 since the Labor Force Surveys have stopped reporting labor market status by marital status.

the labor force. The OECD average at the same point in time was 70.2 percent (OECD, 2009b).9 Male participation rates, on the other hand, are about average: in 2009, the male participation rate (for males aged 25–54) was 93.1 percent which should be compared to an overall OECD average of 92.2 percent.

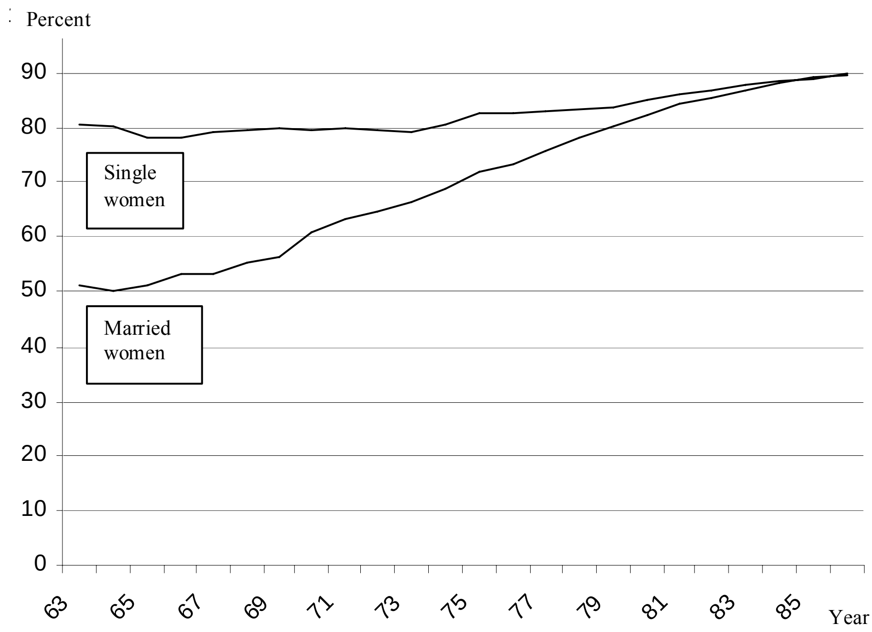

Figure 2 Female labor force participation rates by marital status, 1963-86, percent.

Source: Labor Force Surveys, Statistics Sweden.

2.3 Changes in the composition of children in childcare across cohorts

As explained earlier, childcare was originally targeted at the disadvantaged. But following policy changes in the 1970s, they have become part of an overall policy-package designed to increase (and maintain) female labor force participation rates. One would expect that the increase in labor supply has contributed to change the nature of selection of children into childcare, such that children are increasingly drawn from the higher end of the

[9] This is in line with the cross-country evidence in Jaumotte (2003), which suggests that individual taxation and childcare are two policy tools that contribute to increasing female labor supply. Note also that participation rates among females with small children are higher than among other females, although this to some extent reflects age or cohort effects.

distribution of parental background characteristics starting in 1970. Here we analyze this question in greater detail.

To fulfill this objective we have run earnings regressions using data from LINDA (see Edin and Fredriksson, 2000) in 1970. We conduct this exercise for a single year because it is more convenient to work with a given set of "skill prices"; note, however, that we obtain very similar results if we use additional years.

The LINDA database includes register information on annual earnings, census information on the education level of the subjects, and standard population characteristics derived from the population registers. We restrict attention to males and females, aged 18–59, who are married and have positive earnings. The earnings regression is specified as follows:

$$\ln w_i = \alpha + \gamma ED_i + \beta_1 age_i + \beta_2 age_i^2 + \phi IM_i + \kappa_1 \ln w_i^p + \kappa_2 (\ln w_i^p)^2 + \varepsilon_i \qquad (1)$$

where w denotes annual earnings, γED education level fixed effects, IM controls for immigrant status, and w^p denotes the annual earnings of the partner.[10] Apart from the inclusion of w^p, this is a standard earnings regression. We control for the earnings of the partner since we want to free the other coefficient estimates of the variation in labor supply coming from households with different characteristics.

The estimated coefficients on education, age, and immigrant status are used to predict the earnings of the mothers and fathers in data containing an indicator of whether their child have attended pre-school (these data come from the UGU project; we describe the data in the next section). We think of predicted earnings as a one-dimensional measure of observed skills and use this single index to illustrate how varying labor supply incentives have affected the selection of children into childcare.

Figure 3 illustrates how the changes in labor supply incentives for mothers have affected the composition of children in childcare. It shows how the probability of participating in childcare varies with the potential earnings of the mother (in 1,000s of 1970 SEK) for successive birth cohorts. In the cohort born 1967, there is no (or even negative) selection of children. By the cohort born 1972 – i.e., in just five years – this has changed to positive selection with respect to the earnings potential of the mother. The positive selection becomes even clearer for successive cohorts. Figure 3 thus illustrates that pre-school children become more selected over the time-period spanned by these cohorts.[11]

---

[10] Note that all the results are invariant to estimating the earnings regression in levels.
[11] An analysis along these lines is also presented in Jonsson (2004).

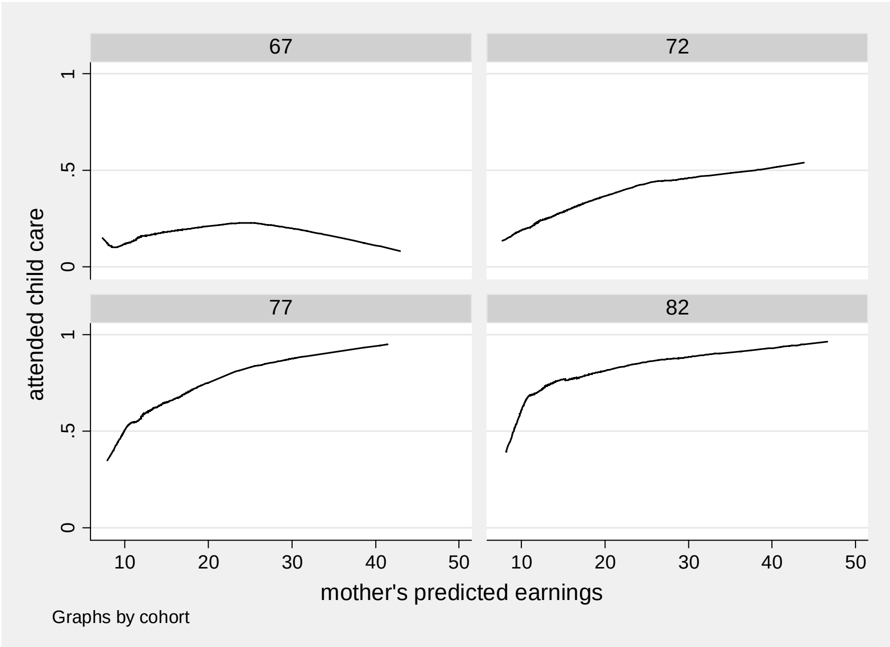

Figure 3 The probability of childcare attendance by mother's skills and cohort

Notes: Predicted earnings in 1,000s of 1970 SEK. Own calculations based on LINDA- and UGU-data as described in the main text. The graph is produced using local linear smoothing.

Figure 4 presents the results of a similar exercise but this time for fathers. The relationship between childcare attendance and the earnings potential of the father is also changing across cohorts. Relative to the mother, the father's earnings potential is not as important in determining childcare attendance of the child. The characteristics of the mother thus appear to be mainly responsible for the changes in the composition of children in childcare that we observe over time.

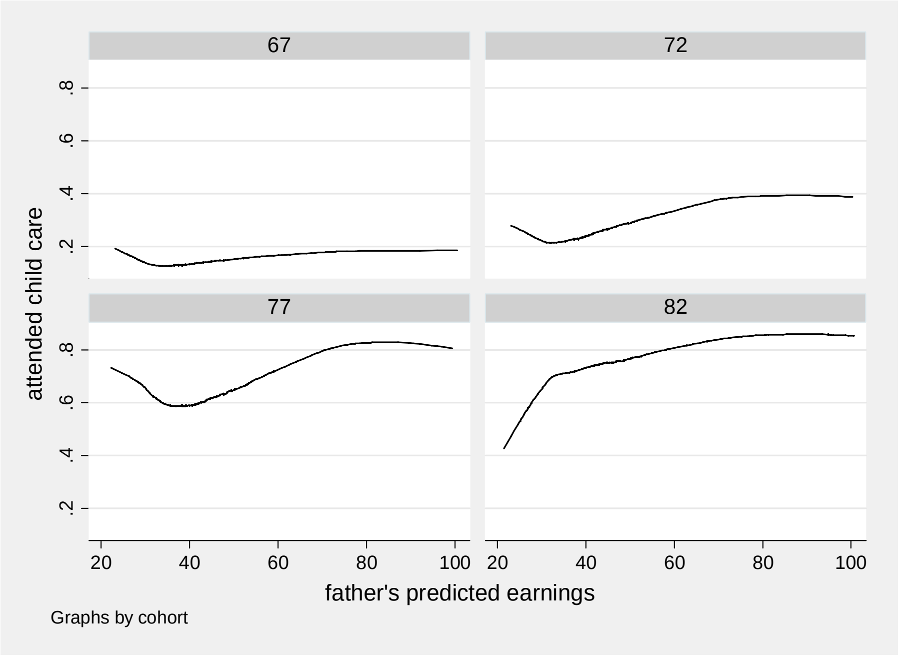

Figure 4 Childcare attendance by father's skills and cohort.

Notes: Predicted earnings in 1,000s of 1970 SEK. Own calculations based on LINDA- and UGU-data as described in the main text. The graph is produced using local linear smoothing.

3 Individual data

We use data from the so-called UGU-project maintained by the Department of Education at Göteborg University; see Härnqvist (2000) for a description of the data. The UGU-data have some features which are very useful for our purposes. Importantly, they include the results of cognitive tests conducted at age 13 for roughly 10 percent of the birth cohorts "born" 1967, 1972, 1977, and 1982.[12] Moreover, the data include information on whether the individuals have attended childcare (pre-schools or family daycare) as well as rudimentary information on how many years they have spent in childcare.

To these data we have matched register information on (individual) educational attainment and information on parental age, education, and immigrant status as well as the number of siblings (in addition we of course have information on the gender and age of the child). The link between parents and child come from the multi-generational register (*Flergenerationsregistret*) which links children to their biological parents.

[12] From 1967 and onwards, the children are sampled in the grade which we would normally expect individuals born a certain year to attend. So, for instance, the "1967-cohort" contains individuals in 6th grade in 1980. Some 95 % of these individuals are actually born in 1967.

The multi-generational register also provides the information on the number of (biological) siblings. Individual and parental education comes from the educational register (*Utbildningsregistret*) which records educational attainment in the Swedish population. Basic demographic information originates from the Population register (*Registret för totalbefolkningen*). These register data are of high-quality; it is unlikely that measurement error is an issue.

Table 1 reports descriptive statistics by childcare status and cohort. Since the data were collected using stratified sampling, we present the weighted means and standard deviations. The descriptive statistics are only reported for the sample which we will use in the regressions. Throughout we condition on the child living in Sweden when they are 24 years old[13]. Moreover, we condition on the there being complete information about the mother. However, we retain observations where the father is either unknown or there is missing information about the educational attainment of the father.

Table 1 indicates that childcare children are favorably selected in terms of their observed characteristics. In particular, the share of mothers with tertiary (compulsory) education is substantially higher (lower) for children who have attended childcare. Between the cohorts born 1967 and 1972 there is a remarkable increase in the share of mothers and (to some extent) fathers with tertiary education who have used childcare, which is much higher than the corresponding increase among parents in general. This is consistent with the view that the tax reform of 1971 improved the incentives for high-skilled mothers to enter the labor market.

Table 1 also illustrates the trend increase in childcare attendance. Between the cohorts born 1967 and 1982, the share who attended pre-schools rises from 15 percent to 76 percent.

[13] The sample is reduced by 216 individuals by conditioning on the individuals being alive and in Sweden at age 24 rather than at age 13. Note that this additional sample reduction has no implications for the effects we estimate on cognitive test outcomes.

Table 1 Descriptive statistics by cohort and pre-school status

	Pre-school status							
	No childcare				Some childcare			
	1967	**1972**	**1977**	**1982**	**1967**	**1972**	**1977**	**1982**
Characteristics of mother								
Compulsory education	0.500	0.364	0.347	0.307	0.312	0.231	0.174	0.133
Upper-secondary education	0.368	0.460	0.510	0.506	0.395	0.379	0.444	0.473
Tertiary education	0.132	0.176	0.143	0.187	0.293	0.390	0.382	0.394
Age at childbirth	26.5	26.5	27.6	28.2	25.5	26.3	27.0	28.1
	(5.5)	(4.9)	(4.8)	(5.1)	(5.4)	(4.6)	(4.7)	(4.9)
Born outside the Nordic countries	0.024	0.023	0.042	0.099	0.022	0.026	0.039	0.061
Characteristics of father								
Missing education	0.068	0.040	0.030	0.051	0.115	0.059	0.039	0.037
Compulsory education	0.420	0.399	0.381	0.322	0.336	0.271	0.237	0.211
Upper-secondary education	0.365	0.395	0.418	0.438	0.353	0.368	0.419	0.436
Tertiary education	0.147	0.166	0.171	0.189	0.196	0.302	0.305	0.316
Age at childbirth*	29.7	29.3	30.5	31.1	28.5	28.8	29.5	30.9
	(6.5)	(5.7)	(5.8)	(5.6)	(6.4)	(5.5)	(5.3)	(5.4)
Born outside the Nordic countries*	0.030	0.034	0.036	0.091	0.074	0.047	0.055	0.072
Father missing	0.011	0.004	0.005	0.013	0.025	0.004	0.008	0.007
Child characteristics								
Female	0.508	0.492	0.479	0.486	0.497	0.490	0.479	0.497
# of siblings (age 12)	1.6	1.5	1.8	2.0	1.3	1.3	1.4	1.6
	(1.0)	(1.0)	(1.1)	(1.2)	(0.9)	(0.9)	(1.0)	(1.0)
Child outcomes								
Language ability (rank)	49.6	48.5	47.5	45.5	52.5	54.8	51.5	51.7
	(28.9)	(28.7)	(29.1)	(28.8)	(28.5)	(28.7)	(28.5)	(28.7)
Inductive ability (rank)	50.0	49.3	46.7	46.3	50.0	52.3	52.1	51.4
	(29.0)	(29.0)	(29.0)	(28.7)	(27.9)	(28.3)	(28.5)	(28.8)
Academic upper-secondary education	0.404	0.454	0.438	0.478	0.484	0.551	0.577	0.600
N	4933	4317	1104	1330	871	1433	1889	4189
(Share of cohort)	(0.85)	(0.75)	(0.37)	(0.24)	(0.15)	(0.25)	(0.63)	(0.76)

Notes: The table reports weighted means and standard deviation, using the sampling probabilities in each strata as weights. * Descriptive statistics are only reported for fathers who are not missing.

The lower half of the table reports the means of our outcome variables – the percentile ranked results on two cognitive tests taken at age 13 as well as the probability of having attained a 3-year "academic" – i.e. university-preparatory – upper-secondary degree (at age 24). The inductive test requires the respondent to fill in the next number in a sequence of numbers. The language test involves finding a word having the opposite meaning as a given word.

We have percentile ranked the results of the cognitive tests within cohort. Across cohorts, the test results fall for children who have not attended childcare and there are corresponding increases for children who have participated in childcare. This pattern may reflect the fact that pre-school children get more favorably selected over time.

Since we focus on the gap between children with an immigrant and native background, it is interesting to examine the immigrant/native gaps by childcare attendance. Throughout the paper, we define immigrant background as both parents being born outside the Nordic countries. Table 2 reports these outcomes.

Table 2 Differences in outcomes by immigrant background and childcare attendance

Childcare	**Immigrant background** (both parents born outside the Nordic countries)		**Difference**
	No	Yes	
	Outcome: Language ability		
No	49.0	23.7	–25.3
			(1.3)
Yes	53.2	32.7	–20.5
			(1.6)
Difference	4.2	9.0	**4.8**
	(0.4)	(2.1)	**(2.1)**
	Outcome: Inductive ability		
No	49.1	37.2	–12.0
			(1.6)
Yes	52.0	42.9	–9.1
			(1.6)
Difference	2.8	5.7	**2.9**
	(0.4)	(2.2)	**(2.2)**
	Outcome: Academic upper-secondary degree		
No	0.436	0.505	0.069
			(0.028)
Yes	0.581	0.619	0.038
			(0.028)
Difference	0.145	0.114	**–0.030**
	(0.007)	(0.039)	**(0.040)**

Notes: Robust standard errors in parentheses. “Difference-in-differences” estimates (bold numbers) are based on regressions including 20,216 observations.

Table 2 conveys several messages. First, individuals with an immigrant background have substantially lower test performance at age 13 than individuals with a native background; the gap in language ability is particularly large. Second, the gaps in cognitive test results between immigrants and natives are smaller among children who have attended childcare; however, it is only the reduction in language ability which is statistically significant. Third, despite the gaps in cognitive test performance, the probability of attaining an academic upper-secondary degree is higher among individuals with an immigrant background; moreover, among the individuals attending childcare the advantage in favour of immigrants is lower than among individuals with no childcare experience, although not significantly so.

Our purpose next is to examine whether these preliminary conclusions hold up to more rigorous analysis.

4 Empirical analysis

The main purpose of this section is to examine how childcare attendance affects future cognitive outcomes and long-run educational attainment. For various reasons we will rely on a selection of observables assumption. The main reason for making this assumption is that we have found no credible instrument which can be used to estimate the effects of interest. Furthermore, the virtue of an instrument is not all that obvious in this case. Since the nature of the selection differs across cohorts, a valid instrument will most likely yield different estimates across cohorts just because the set of "compliers" vary across cohorts (see Imbens and Angrist, 1994).

Our main approach for examining whether selection on observables is a reasonable assumption is to vary the set of conditioning variables. If the coefficients of interest do not vary with the conditioning set we view the results as being robust.

To preview our results, we conclude that we cannot credibly estimate the average effect of childcare attendance. However, we consistently find that childcare attendance reduces the gap in language skills by immigrant background.

4.1 Empirical set-up

We specify the outcome equations as follows

$$y_{ijc} = \alpha_j + \alpha_c + \beta CC_{ijc} + \gamma(s \times CC)_{ijc} + \lambda s_{ijc} + \varphi_1 X_{1,ijc} + \varphi_2 X_{2,ijc} + \varepsilon_{ijc} \qquad (2)$$

where i indexes individuals, j municipalities, and c cohorts; thus α_j (α_c) denotes a municipality (cohort) fixed effect.

y_{ijc} denotes the outcome of interest, i.e., either the percentile ranked results on the (two) cognitive tests or educational attainment. The tests were conducted in 6th grade, when the children were aged 13. Educational attainment is measured at age 24; we specify this outcome as the probability of having at least 3 years of university-preparatory upper-secondary education.

CC_{ijc}, the treatment of interest, is defined to equal unity if the parents respond that their child has attended childcare; it equals zero if the parents have responded not at all. We also interact the treatment with indicators of the family background of the children (*s*). We consider two such interactions. We estimate a separate effect for children: (i) whose parents are both born outside the Nordic countries; and (ii) whose mother has only compulsory education.

We will also examine whether there are differential effects across childcare modes. Thus we define separate indicator variables for children who have attended pre-schools and family daycare and interact these alternative treatment indicators with the family background of the children.

The two last pieces of notation in equation (2) concern the control variables that we include in the regression. The first set of variables ($X_{1,ijc}$) includes predetermined characteristics which should be included to control for selection on observed characteristics. The variables included in $X_{1,ijc}$ are basically the ones listed in Table 1. The other set of variables ($X_{2,ijc}$) include the variables that we will use to "test" our selection-on-observables assumption. The underlying idea is that if the estimates are plagued by selection (or omitted variables) bias and if the inclusion of $X_{2,ijc}$ moderates (or eliminates) this bias we should see substantial changes in the coefficients of interest when we control for $X_{2,ijc}$. In practice, this idea has been around for quite some time; Altonji et al. (2005) provides a formal justification for such sensitivity analyses.

In the current application, we will include the result on a spatial ability test. The inclusion of this test arguably controls for selection. The problem with including it is that the test is conducted at age 13. Therefore, the variation in spatial ability is potentially an outcome of childcare attendance. However, evidence reported by Cahan and Cohen (1989) as well as the recent evidence presented in Öckert (2009) suggests that spatial ability is less malleable to schooling than inductive and language ability. According to Öckert, a year of schooling improves inductive and verbal ability by 0.17–0.18 standard deviations, but spatial ability "only" increases by 0.07 standard deviations.

4.2 The distributional impact of childcare

Table 3 presents the results. Columns (1)–(3) report the results for language ability, while columns (4)–(6) contain the results for inductive ability ("number series"). Panels A–B consider the interaction between childcare attendance and immigrant background (panel A) and mother's education (panel B), respectively.

Looking at Table 3 it is clear that the main effect of childcare attendance is not credibly identified. While the correlations presented in columns (1) and (4) are all positive and significant, they are all rendered insignificant just by controlling for observed characteristics; see columns (2) and (5). If selection on observed and unobserved characteristics works much in the same way it is not hard to imagine that the main effects would be reduced further. The evidence reported in columns (3) and (6) is consistent with this conjecture. Here we include the measure of spatial ability which reduces the size of the main effect further.

Under the assumption that selection is the same across groups (we will relax this assumption below), it may still be meaningful to examine the distributional impact of childcare attendance. The second row in each panel thus contains the estimates of the interaction between childcare attendance and immigrant background and the mother being less educated, respectively. It seems that childcare attendance narrows the distribution of language skills for children with different immigrant backgrounds. The effects on language skills do not vary by mother's education, however. Moreover, there is no distributional impact of childcare attendance on inductive skills.

It is noteworthy that the interaction estimate for children with immigrant background stays almost the same when we control for spatial ability. This suggests that selection is not driving the interaction estimate.

What does the estimate on the interaction between childcare and immigrant background imply? Jonsson (2004) shows that, on average, the children with some childcare experience in the cohorts born 1966–81 have spent roughly three years in childcare. The estimate in column (3) thus suggests that each year of childcare experience reduces the gap in language ability between immigrants and natives by 2.7 (8.1/3 = 2.7) percentile ranks. The raw gap between immigrants and natives with no childcare experience amounts to 25 percentile ranks (see Table 2 or column (1), panel A). Thus, each year of childcare experience closes 10 percent of the gap between immigrants and natives; 5 years of pre-school experience reduces the gap by 50 percent. These effects are rather substantial, suggesting that childcare is an important vehicle for closing the gap between immigrants and natives in terms of language ability.

Table 3 Effects of childcare attendance on cognitive outcomes by family background

	Language ability			Inductive ability		
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Immigrant background (both parents born outside the Nordic countries)			
Childcare, main effect	4.25	0.645	0.319	3.36	0.458	0.048
	(0.63)	(0.620)	(0.584)	(0.63)	(0.618)	(0.564)
Childcare interaction	7.92	9.04	8.11	0.488	0.817	–0.091
	(2.57)	(2.44)	(2.42)	(2.580)	(2.547)	(2.34)
Main effect	–25.0	–20.7	–18.8	–11.9	–14.0	–11.5
(immigrant background)	(1.7)	(2.6)	(2.5)	(1.9)	(2.9)	(2.8)
Spatial ability	--	--	0.344	--	--	0.431
			(0.008)			(0.008)
Cohort FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Basic covariates	--	Yes	Yes	--	Yes	Yes
Adjusted R^2 (within municipality)	0.025	0.126	0.233	0.015	0.073	0.250
N	20,126	20,126	20,126	20,126	20,126	20,126
			B. Mother less educated (no more than compulsory education)			
Childcare, main effect	4.41	1.34	0.808	3.40	1.13	0.452
	(0.69)	(0.68)	(0.634)	(0.69)	(0.68)	(0.612)
Childcare interaction	–4.66	–1.68	–0.978	–4.64	–2.46	–1.58
	(1.22)	(1.21)	(1.139)	(1.22)	(1.22)	(1.11)
Main effect	–9.13	–12.9	–10.1	–7.52	–10.6	–7.04
(low education)	(0.68)	(0.8)	(0.8)	(0.69)	(0.8)	(0.77)
Spatial ability	--	--	0.344	--	--	0.431
			(0.008)			(0.008)
Cohort FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Basic covariates	--	Yes	Yes	--	Yes	Yes
Adjusted R^2 (within municipality)	0.040	0.120	0.232	0.031	0.073	0.250
N	20,126	20,126	20,126	20,126	20,126	20,126

Notes: Linear regression models estimated using the inverse sampling probabilities as weights. Robust standard errors in parentheses. Basic covariates include: gender; number of siblings; the mother's educational attainment, age at childbirth, and immigrant background; the father's educational attainment, age at childbirth, immigrant background, and an indicator for unknown father; the immigrant status of both parents.

Next, let us turn to the effects on long-run educational attainment. Constraints related to data quality force us to focus on the probability of having at least a 3-year university-preparatory – an “academic” – upper-secondary degree. We measure this outcome at age 24.[14]

Table 4 presents the results. Despite the fact that childcare improves the language ability of immigrants relative to natives, there is no differential effect on the probability of attaining an academic upper-secondary degree. Moreover, there is no differential effect by mother’s education (see panel B), which is consistent with there being no differential effects of childcare attendance on the cognitive outcomes by mother’s education.

Determining exactly why there are no differential effects by immigrant background is, to some extent, a matter of speculation. But it seems that the effect on language skills is too small to alter the choices made by children with an immigrant background. Note, in this respect, that the main effect of having an immigrant background is consistently positive, despite the fact that immigrants have both lower test results at age 13 and lower grade point average (GPA) when leaving compulsory school. Thus, cognitive skills (as measured by the tests or GPA) have a smaller impact on subsequent educational choices among immigrants than among natives.

Table 4 Effects of childcare attendance on educational attainment (at least academic upper-secondary degree at age 24) by family background

	(1)	(2)	(3)
		A. Immigrant background (both parents born outside the Nordic countries)	
Childcare, main effect	0.097	0.020	0.017
	(.011)	(0.011)	(0.010)
Childcare interaction	0.005	0.021	0.012
	(.051)	(0.048)	(0.048)
Main effect (immigrant background)	0.055	0.082	0.101
	(0.036)	(0.053)	(.054)
(Spatial ability)/100	--	--	0.337
			(0.014)
Cohort FE:s	Yes	Yes	Yes
Municipality FE:s	Yes	Yes	Yes
Basic covariates	--	Yes	Yes
Adjusted R^2 (within municipality)	0.032	0.160	0.196
N	20,126	20,126	20,126

[14] High-quality information on educational attainment is available to us 1991–2006. In 1991, the oldest cohort (those born 1967) are 24 years-old, while, in 2006, the youngest cohort (those born in 1982) is 24 years-old. It would have been preferable to record educational attainment at a higher age, because then we could have included tertiary education. Alternatively, a more “discriminatory” outcome would be the probability of having an upper-secondary degree at age 19 (which is the normal graduation age). However, none of these two options are open to us because of data constraints.

	B. Mother less educated (no more than compulsory education)		
Childcare, main effect	0.077 (.012)	0.023 (0.012)	0.018 (0.011)
Childcare interaction	–0.055 (0.021)	–0.010 (0.020)	–0.003 (0.020)
Main effect (low education)	–0.203 (0.012)	–0.286 (0.014)	–0.258 (0.014)
(Spatial ability)/100	--	--	0.337 (0.014)
Cohort FE:s	Yes	Yes	Yes
Municipality FE:s	Yes	Yes	Yes
Basic covariates	--	Yes	Yes
Adjusted R^2 (within municipality)	0.071	0.160	0.196
N	20,126	20,126	20,126

Notes: Linear probability models estimated using the inverse sampling probabilities as weights.. Robust standard errors in parentheses. Basic covariates include: gender; number of siblings; the mother's educational attainment, age at childbirth, and immigrant background; the father's educational attainment, age at childbirth, immigrant background, and an indicator for unknown father; the immigrant status of both parents.

4.3 Robustness checks and extensions

The purpose of this section is to present some robustness checks and extensions of our baseline specification. Throughout, we focus on the differential effect on language ability by immigrant background. We view the estimate presented in column (3) of Table 3 as our baseline result. In Table 5 we examine whether this estimate is robust to alternative assumptions and alternative definitions of treatment; for convenience the first row reproduces the baseline estimate.

An identifying assumption underlying our baseline results is that the process determining selection into childcare is the same on average for immigrants and natives. To be more precise, we assume that the correlations of the unobserved and observed variables (e.g. ability) are linear in the covariates and the same across the two groups. In model (2) we relax these assumptions by allowing the coefficient on spatial ability to vary by immigrant background; moreover, we introduce the predicted earnings of the mother which is allowed to have a separate effect for children with an immigrant background.[15] If the estimates are unaffected by these extensions

[15] The reason for introducing the linear earnings index, rather than interacting observed covariates fully with immigrant status, is that we want to save on degrees of freedom. To obtain the earnings predictions we have estimated equation (1) separately for immigrant and native mothers. Note that it does not matter for the results if we use a single set of estimates for both groups.

we interpret this as saying that (potential) unobserved variables varying across the two groups does not bias the baseline estimates.

Row (2) shows that allowing for differential selection by immigrant background has no implications for the baseline result (if anything the result is strengthened). The estimate on the interaction between immigrant background and childcare attendance equals 9.2 with a t-ratio of 3.8.

The model in (3) includes separate treatment effects for pre-schools and family daycare. As explained earlier these two childcare modes imply different kinds of treatments. The pedagogical content may be higher in pre-schools but group sizes are also higher. The magnitudes of the estimates imply that family daycare reduces the gap between immigrants and natives more than pre-schools, although the estimates are not different from each other in the statistical sense.

Finally, in (4) we estimate the model in (3) separately by cohort. There is some variation in the estimate of the treatment interactions across cohorts. The lower bounds of the (95 percent) confidence bands estimated for the 1967 cohort are higher than the point estimate of the pooled regression in model (3). The confidence bands of all other cohorts cover the corresponding estimate of model (3). It is also noteworthy that family daycare consistently appears to reduce the language gap more than pre-schools.

In sum, we view the variations reported in Table 5 as lending support to the baseline estimates reported in panel A) of Table 3. Childcare attendance thus reduces the gap in language skills between immigrant and native children.

Table 5 Effect of childcare attendance on language ability by immigrant background, variations of the baseline specifications

		Main effect	Interaction with immigrant background	Observations	Adj. R^2
(1):	**Baseline estimate**	0.319	8.11	20,126	0.233
		(0.584)	(2.42)		
(2):	**Allowing for differential selection by immigrant background**	0.226	9.16	20,126	0.237
		(0.583)	(2.39)		
(3):	**(2) with separate treatment effects for pre-school and family daycare**			20,126	0.236
	Pre-school	0.282	5.72		
		(0.579)	(2.52)		
	Family daycare	–0.201	10.5		
		(0.676)	(4.9)		
(4):	**(3) estimated by cohort**				
	1967			5,864	0.237
	Pre-school	–0.689	21.3		
		(1.354)	(6.3)		
	Family daycare	1.87	28.3		
		(1.89)	(3.9)		
	1972			5,750	0.222
	Pre-school	1.26	0.290		
		(1.19)	(7.382)		
	Family daycare	3.72	22.6		
		(2.05)	(18.4)		
	1977			2,993	0.253
	Pre-school	–0.400	9.00		
		(1.222)	(5.83)		
	Family daycare	–0.669	13.1		
		(1.223)	(9.2)		
	1982			5,519	0.260
	Pre-schools	0.502	7.22		
		(0.916)	(3.05)		
	Family daycare	–1.09	10.0		
		(0.92)	(5.1)		

Notes: Model (2) includes an interaction between immigrant background and spatial ability, the predicted earnings of the mother, and the predicted earnings of the mother interacted with immigrant background. The models in (4) are estimated separately by cohort. Robust standard errors in parentheses. Regressions are weighted using the inverse of the sampling probabilities. Covariates include: gender; number of siblings; the mother's educational attainment, age at childbirth, and immigrant background; the father's educational attainment, age at childbirth, immigrant background, and an indicator for unknown father; the immigrant status of both parents; and spatial ability.

4.4 Summary and discussion of the results

Let us summarize the results. Comparing children of immigrant and native background we find that:
a) the gap in language skills across the two groups is reduced
b) family day care appears to reduce the gap more than regular pre-schools
c) there is no effect on the gap in inductive skills across these two groups
Comparing children with low and high-educated mothers we find that
d) there are no effects on the gaps in language and inductive skills

What is the rationale for these (medium-run) findings? We think that the configuration of the results suggests that what childcare offers is mainly an arena for interaction with other children as well as staff. Any pedagogical treatment effects appear to be limited – or at least not substantial enough to alter medium-run cognitive achievement.

If pedagogical content would have been an important part of the treatment we would expect to see a reduction in the gap in inductive skills as well, a decrease in the cognitive ability gaps along the educational dimension, and regular pre-schools to have greater effects on the achievement gaps than family daycare.

Rather we observe: a reduction in the language ability gap only; this effect only shows up along the immigrant/native comparison; and, if anything, family daycare has a greater effect on the language ability gap among immigrants and natives. This suggests to us that childcare furthers the language ability of immigrants since it opens up for closer interaction with native-born children and Swedish speaking staff.

We have not been able to detect any differential long-run effect of childcare on educational attainment. It may be that the effect on language skills is too small to alter the choices made by children with an immigrant background.

5 Concluding remarks

In this paper we have estimated the relationship between childcare attendance and medium and long run educational outcomes. We have done this using data on individuals born between 1967 and 1982.

The time period spanned by these cohorts featured a substantial expansion of childcare: in 1975, 18 percent of children aged 1–5 attended childcare; by 1985 (in just 10 years) the share of 1–5 year-olds participating in childcare had increased to 56 percent. The childcare expansion was intimately tied to the increase of the labor force participation of women. We have illustrated that, across cohorts, children in childcare were increasingly drawn from the higher end of the distribution of family background characteristics.

The changes in the composition of participating children raise issues regarding the selection into childcare. For that reason we have focused on whether childcare attendance has differential effects by immigrant background. We have found that childcare participation narrows the language ability gap between children with an immigrant background and children with a native background. Our estimates imply that a year of childcare experience reduces the overall gap between immigrants and natives in language ability by 10 percent. This conclusion is robust to allowing differential selection across immigrants and natives.

We have found no differential effects on inductive skills, however. Nor does childcare affect the distribution of longer-run educational attainment. The latter result is somewhat surprising, given that the gap in language skills is affected by childcare attendance. Taken seriously, it is perhaps due to the effect on language skills being too small to alter the educational choices of immigrants; educational choices of immigrants in Sweden seem to be driven by cognitive ability to a lesser extent than among natives. But for (at least) two reasons it would be premature to conclude that childcare has no differential long-run effects. First, since some of the individuals included in the analysis are born in the 1980s, we measure educational attainment at a relatively young age (24 years-of-age). Therefore, we have focused on the probability of attaining a university-preparatory upper-secondary degree. Since we cannot account for tertiary education we may miss some of the potential effect on educational attainment. Second, we are perhaps ultimately interested in whether the differential effects on language ability feed on to long-run earnings outcomes. Again, the time-span of our data precludes such an analysis.

In contrast to the vast majority of U.S. states, childcare in Sweden are not targeted at the disadvantaged. Rather it is universally available; during the time period we have considered it was in fact targeted at the employed. Disadvantaged (particularly immigrant) children are less likely to participate in pre-school education. The evidence we have offered suggests that increasing the childcare participation rates among immigrant children will close some of the gap between natives and immigrants in language skills.

References

Altonji, J., T. Elder, and C. Taber (2005): Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, Journal of Political Economy, vol. 113, no. 1

Baker M., J. Gruber, and K. Milligan (2008): Universal childcare, maternal labor supply and family well-being, Journal of Political Economy, vol. 116, no. 4

Berlinski, S., S. Galiani, and P. Gertler (2009): The effect of pre-primary education on primary school performance, Journal of Public Economics , vol. 93, no.1-2

Cahan, S. and N Cohen (1989): Age versus schooling effects on intelligence development, Child Development, vol. 60, no. 5
Cuhna, F., J. Heckman, L. Lochner, and D. Masterov (2006): Interpreting the evidence on life cycle skill formation, in E. Hanushek and F. Welch (eds.), Handbook of the Economics of Education, North-Holland, Amsterdam, 697-812.
Datta Gupta, N. and M. Simonsen (2007): Non-cognitive child outcomes and universal high quality child care, IZA Discussion Paper no. 3188.
Edin, P-A. and P. Fredriksson (2000): LINDA – Longitudinal INdividual DAta for Sweden, Working Paper 2000:19, Department of Economics, Uppsala University.
Gormley, Jr., W. and T. Gayer, (2005): Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program, Journal of Human Resources, vol. 40, no. 3
Härnqvist, K. (2000): Evaluation through follow-up, in C-G. Jansson (ed.), Seven Swedish Longitudinal Studies in the Behavioral Sciences, Forskningsrådsnämnden, Stockholm.
Havnes, T. and M. Mogstad (2009): No child left behind. Universal child care and children's long-run outcomes, Discussion Papers No. 582, Statistics Norway.
Imbens, G. and J. Angrist (1994): Identification and estimation of local average treatment effects, Econometrica, vol. 62, no. 2
Jaumotte, F. (2003): Female labour force participation: past trends and main determinants in OECD countries, Working Paper 376, OECD Economics Department.
Johansson, G. and I-B. Åstedt (1996): Förskolans utveckling – fakta och funderingar, HLS Förlag, Stockholm.
Jonsson J. (2004): Förskola för förfördelade, in Bygren, M et al. (eds.), Familj och arbete – vardagsliv i förändring. SNS förlag, Stockholm.
Karoly, L., M. Kilburn, and J. Cannon (2005): Early childhood interventions: Proven results, future promise, RAND Corporation, Santa Monica.
Magnuson, K., C. Ruhm, and J. Waldfogel (2007): Does prekindergarten improve school preparation and performance? Economics of Education Review, vol. 26, no. 1
Öckert, B. (2009): Does going to high school make you smarter?, manuscript, IFAU.
OECD (2001): Starting strong. Early childhood education and care, Organization for Economic Co-operation and Development, Paris.
OECD (2006a): Where Immigrants Succeed: A Comparative Review of Performance and Engagement in PISA 2003, Organization for Economic Cooperation and Development.
OECD (2006b): Starting strong II. Early childhood education and care, Organization for Economic Co-operation and Development, Paris.
OECD (2009a): Education at a glance 2009, Organization for Economic Co-operation and Development, Paris.
OECD (2009b): Employment outlook, Organization for Economic Co-operation and Development, Paris.
Richardson, G. (2004): Svensk utbildningshistoria, Studentlitteratur, Stockholm.
Schneeweis, N. (2009): Educational institutions and the integration of migrants, forthcoming, Journal of Population Economics.
Selin, H. (2008): Four empirical essays on responses to income taxation, Economic Studies 110, Department of Economics, Uppsala University.
SOU 1944:20: Skolan i samhällets tjänst, 1940 års skolutrednings betänkanden I.

Essay 3: Gender bias in a gender equal society – evidence from Swedish parental leave use[1]

1 Introduction

Early childhood experiences and interventions are important determinants of later-life performance and it is obvious that not only schools, but also family environment matter (Cunha and Heckman, 2007; Doyle et al, 2009). It has for example been found that maternal employment when the child is very young is detrimental for the child's development (see for example Berger et al, 2005; Bernal, 2008; Blau and Grossberg, 1992, Desai et al, 1989; Ruhm, 2000 and 2004; Tanaka, 2005) and that mothers and fathers may affect child development differently (Averett et al, 2007, Parcel and Menaghan, 1994). At the same time, there is growing evidence on differential treatment of sons and daughters also among industrialized countries, and this may in turn affect child outcomes. (In addition, if parental behaviour differs depending on child gender, this is of academic interest since it questions the use of child gender as an exogenous instrument in some applications, see for example Bennedsen et al, 2007, or Chun and Oh, 2002).

The purpose of this paper is to investigate whether Swedish parents treat sons and daughters differently with respect to parental leave. Does the gender of the child affect the probability and length of parents' parental leave? Child gender is regarded exogenous, and this is possible since we focus on the first-born child (the gender of later-born siblings may not be treated as exogenous; see below). A large, register-based and population-wide data set virtually free from attrition and missing variables enables precise estimation of the effect also among different subgroups of families. By focusing on Sweden, the most gender equal country in the world as measured by the United Nations gender-related development index (GDI) and gender empowerment measure (GEM) (United Nations, 1995), yields additional evidence on parental gender biases in the industrialized world. The paper also addresses the issue of who is driving the bias – mothers or

[1] I am grateful to Peter Fredriksson and Per Johansson for valuable guidance. I would also like to thank Mikael Elinder, Jenny Nykvist, Håkan Selin, Peter Skogman Thoursie and Björn Öckert, as well as seminar participants at the Department of Economics, Uppsala university, for important suggestions and comments. The financial support from the Swedish Council for Working Life and Social Research, FAS (dnr 2004-1222) is also acknowledged.

fathers or both? – by using proxies for intra-family bargaining power. In addition, data on temporary parental leave (for sick children) is used to examine the competing hypothesis of gender differences in sickness as the sole explanation for any differential treatment depending on gender.

Previous studies on gender biases in industrialized countries are primarily based on US data. They show that gender biases are prevalent not only in the developing world (as shown by skewed sex-rations and "missing girls") but also in many industrialized countries, they are only expressed differently (see Lundberg, 2005a, for an overview). For example, a first-born boy is more likely to live with the father because a boy increases the likelihood that the parents marry, decreases the risk of divorce, and increases the likelihood of paternal custody in case of divorce (Dahl and Moretti, 2008, Lundberg 2005a, Lundberg et al, 2007). Parents of boys also report greater marital satisfaction and fathers spend more time with sons (in contrast to mothers, who often are found to be relatively gender neutral in this respect) (Lundberg 2005a). Sons also affect parental work hours although the direction of the effect may differ between studies (Lundberg, 2005a, b). Some studies find that mothers allocate relatively more resources to daughters and fathers more to sons (Thomas, 1994) but in general, little evidence is found of differential treatment of children in terms of resource allocation (Lundberg 2005a). Child gender also affects fertility: among married American couples, a first-born boy reduces fertility (Dahl and Moretti, 2008) while evidence from the Nordic countries shows that boys instead increase fertility among mothers (married and unmarried) in Sweden, Denmark and Norway while fertility for Finnish mothers are reduced by first-born boys (Andersson et al, 2004). Stated preference surveys also show relatively large preferences for sons among American fathers but only a slight preference for girls among mothers (Dahl and Moretti, 2008). Also, there is a growing literature on how mothers and fathers, as well as maternal and paternal grandmothers and grandfathers, may affect children differently and that child gender may play a role (Cox, 2007).

This study adds to previous literature by focusing on a new outcome variable (parental leave) in a country with top ratings on gender equality (Sweden). Also, we depart from previous studies in that we utilize proxies for intra-family bargaining-power to locate the origin of any gender bias. Also, the estimates on temporary parental leave helps rule out the competing explanation of sickness differences as the sole explanation for any gender bias.

The results show that child gender has no effect on the parents' probability of taking parental leave. However, among parents who use at least one day of leave, there is a statistically significant effect of child gender on the duration of the leave. A first-born boy increases fathers' parental leave by around 0.6 days and decreases mothers leave by a similar amount, leaving the total time on parental leave virtually unchanged. From a policy perspective, the effect is very small and unlikely to affect child well-being. At the same time, compared to the mean time on parental leave for fathers of girls (42 days) a son increases fathers' time on parental leave by 1.5 percent. This is in line with previous studies from other industrialized countries, both in terms of direction and effect size. There are also interesting differences between groups. The effect of child gender on parental leave is larger among non-traditional families with higher maternal than paternal earnings and/or educational levels. Assuming that relative income or relative educational attainment are reasonable proxies for bargaining power within the family, this indicates that it may be the mothers, rather than the fathers, that are the driving force behind these gender biases.

Finally, it is also shown that these gender biases persist also when looking at temporary parental leave (for older, sick children) when children with exactly the same number of sick days are compared. Hence, gender differences in sickness are not enough to explain the differential treatment of sons and daughters.

2 The Swedish parental leave system

The Swedish parental leave legislation gives the parents 450-480 days of governmentally paid parental cash benefits for each child. Both parents have equal rights to use the system and each parent is assigned half of the days. However, the parents can freely transfer days to each other except from the so called "daddy months". For children born from the 1st of January 1995, 30 days of the parental leave benefit is earmarked for each parent. For children born 1st of January 2002 or later, 60 days are set aside for each parent. There is great flexibility in how the cash benefits can be used, and parents can for example extend their number of days on cash benefits by accepting a lower income replacement for each day. The parental cash benefits can be used until the child turns eight years old.

The reimbursement level of the parental cash benefit varies, both depending on the length of the leave and on the child's birth date. Most days are reimbursed as 80 percent of the previous salary up to a certain ceiling that varies from year to year, while the remaining 90 days is reimbursed on a low flat rate, independent on income. For parents who lack previous income, all days are reimbursed at a flat rate. Fore more details about the number of parental leave days and the reimbursement levels over the years, see Section A2 in Appendix.

In addition to the parental leave benefit, there is also a *temporary* parental leave benefit. It is similar to the parental leave benefit, but focuses on the parents' right to stay home from work to care for sick children. The temporary parental benefit is aimed at parents who work, and hence it is not available for parents on parental leave (with a few exceptions, mainly regarding children in need of hospital treatment). In other words, parents currently on parental leave can not exchange their parental leave days for temporary parental leave. This means in practise that temporary parental leave mostly is used for older children.

Although women and men now have the same rights in the parental leave system, large differences remain in how they use the system. Despite a dramatic rise in fathers' share of days with parental leave benefits – from 9.9 percent in 1997 to 17.5 percent in February 2004 – 15 percent of the fathers use no parental leave benefits at all, and 25 percent do not use their ten paternal days (Batljan et al, 2004). There are large differences between groups in this respect. Immigrant fathers tend to use the parental leave system less on average, while more educated fathers and fathers with higher incomes use it more (The Swedish Social Insurance Agency, 2005). The parental leave system is also used differently by mothers and fathers. For mothers, there is a weaker link between time absent from work to take care of the child and parental leave days with cash benefits, since many mothers extend their time home with the child by accepting a lower income replacement for each day. In contrast, fathers' days on parental leave is strongly correlated with his days on cash benefits (The Swedish Social Insurance Agency, 2004a). In addition, mothers and fathers use the parental leave benefits at different child ages. Before the child turns one year old, the mothers are usually the ones staying home with the child, while fathers' use of parental leave benefits is highest when the child is around one year old (ibid). For a more detailed description of the Swedish parental leave system and trends in the benefit usage, see for example The Swedish Social Insurance Agency (2004a; 2004b; 2005).

3 Data, estimation and identification strategy

3.1 Data and estimation strategy

The data used combine register information from the LISA data base and the so called multigenerational registry, provided by Statistics Sweden, with parental leave data from the Swedish Social Insurance Agency. It encompasses the entire Swedish population and contain high-quality, individual level information on all children and their family members with virtually no attrition or missing variables problems. In addition to information on parental leave, a number of control variables are also available including age, marital status, earnings and educational levels of the parents.

Our sample consists of all native Swedish families[2] whose first child was born 1993-2005. In the main estimations, the focus is on parental leave during the first 24 months of the child's life. (Focusing on early parental leave when most families have not yet received additional children isolates the effect of the first-borns gender; longer run parental leave may also be affected by the younger siblings' gender). In the extended analysis, this restriction is relaxed and also the effect on longer run parental leave is investigated.

Table 1 shows some descriptive statistics for the sample, by child gender. All variables are measured before the birth of the child. Clearly, there are no differences in terms of parental characteristics between families whose first-born is a daughter compared to a son. All coefficients are similar in magnitude for boy and girl-families and there are no statistically significant differences at conventional levels, alfa<=10%. Fathers are older and have higher pre-birth earnings than mothers, but a larger proportion of the mothers have a university degree. Relatively few (<20 percent) are married which is due to the fact that we measure marital status on average one year before the birth of the child.

Table 1 Descriptive statistics, by child gender

	First-born daughter	**First-born son**
Father's age	28.56	28.57
	(4.954)	(4.933)
Mother's age	26.52	26.52
	(4.560)	(4.546)
Father w. high school educ.	0.584	0.586
	(0.493)	(0.493)
Mother w. high school educ.	0.538	0.537
	(0.499)	(0.499)
Father w. university educ.	0.301	0.301
	(0.459)	(0.459)
Mother w. university educ.	0.363	0.362
	(0.481)	(0.481)
Father's earnings	194.1	194.4
	(148.3)	(158.5)
Mother's earnings	148.9	148.8
	(102.3)	(98.87)
Married	0.195	0.194
	(0.396)	(0.396)
N	173423	183348

Notes: There are no statistically significant differences between the groups at conventional levels (10% or below). Standard errors in parentheses.

[2] Among immigrants, there are missing information primarily on educational levels for around 20 percent of all individuals.

The analysis focuses on the effect of the first-born child's gender on 1) the probability of taking parental leave, and 2) the duration of parental leave, given duration >0. This is done separately for mothers and fathers. Although it is possible that each spouse's parental leave decision is affected by the other spouse's behavior, we can still identify the effect of child gender for each parent separately (since child gender is exogenous). What we cannot separate is what mechanisms are at play: the direct effect of child gender on each parent's leave and/or the indirect effect of child gender via spousal responses (for example gender neutral mothers who change their parental leave as a response of changed paternal behavior). Hence, we cannot identify whether it is the father, the mother, or both that are gender biased.

The models are estimated using ordinary least squares (OLS). If the gender of the first-born child is exogenous, no control variables are necessary to identify the causal effect. However, in order to possibly increase precision and as an informal way of testing the exogeneity assumption, I include the standard control variables mentioned above.

3.2 Identification: is child gender exogenous?

To identify causal effects and not only correlations, child gender is required to be exogenous (at least given the covariates used). There are two main issues regarding exogeneity of child gender. First, biological research has shown that there are, in fact, small correlations between different characteristics of the parents and the gender of the child. Second, if there are selective abortions depending on the gender of the child, then the exogeneity-assumption is clearly violated.

First, consider the biological differences. For long, it has been known that slightly more boys than girls are being born. The normal sex-ratio lies in the range of 104-107 boys for every 100 girls (Chahnazarian, 1988) and this is believed to be natures way of compensating for a higher mortality among males (The National Board of Health and Welfare, 2002).

There is also a growing literature showing that there are differences in the likelihood of having boys versus girls between different individuals and individuals living in different environments. According to the Trivers-Willard-hypothesis (Trivers and Willard, 1973 and Wells, 2000) mothers in good condition should have more sons than mothers in poorer condition. The mechanism is a higher male mortality and morbidity in general, which in turn is more dangerous when the mother is of poorer condition. The explanation offered is evolutionary – while almost all girls conceive and get an offspring, only males in good condition (born by mothers in good condition) are able to mate and have an offspring (see the papers for further details). This theory has also been supported empirically although the definition of what constitutes a woman in "good condition" varies. For example, Almond and Edlund (2006) find that married and better educated women bear more sons, while Cagnacci et. al (2004) find that women with low pre-pregnancy weight and a greater weight gain during pregnancy bear fewer sons. Chahnazarian (1988) reviews the literature on the determinants of child gender and finds both that blacks have a larger proportion of girls than whites, and that parents with higher socioeconomic status are more likely to have sons. But he concludes that all the effects found are small. There is also evidence that a stressful environment reduces the number of boys being born. Catalano et al (2005a) show that the share of boys born in California dropped shortly after the September 11 terrorist attacks. In another paper, Catalano et al (2005b) show that male fetal deaths increase in times of high unemployment rates.

So, there seem to be evidence of biological and social determinants of child gender. Is this enough to bias my results? I argue this is unlikely, for several reasons. First, the effects of different social characteristics on a child's gender are usually small in magnitude. For example, in Almond and Edlund (2006) a mother without a high school degree is 0.3 percent less likely to have a boy compared to a more educated mother. Since our estimates of child gender bias for fathers are much higher (1.5 percent; see Table 2 below) it seems unlikely that differences in parental characteristics between parents with sons versus parents with daughters are the sole explanation for why fathers are home more with sons. Second, I do control for many of the social characteristics, such as parental education levels, that has been found to affect child gender. Third, when comparing the characteristics of parents of boys versus parents of girls in my sample, I find no differences at all (see Table 1 above).

Second, consider the issue of selective abortions. From 1975, when the new abortion legislation was introduced in Sweden, it is up to each woman to freely decide about abortions until the 18th week of pregnancy. After that, abortions are only allowed after permission from The National Board of Health and Welfare, and only in the case of severe indications. Since it is possible to detect the child's gender before the end of the 18th week, either by an ultrasound test or by amniocentesis, it is in principle possible for Swedish women to decide to have an abortion after knowing the child's gender[3].

However, descriptive statistics suggests this is not a problem for my study. Prior to 1975, abortions were not free but instead only allowed after special permission. This implies that sex-selective abortions were in practice not possible in the early 1970:s (The National Board of Health and Welfare, 2006). In addition, the techniques for testing child gender have gradually been improved, and are more reliable today than they were some decades ago. Still, when comparing the sex-ratio of boys versus girls over time, there

[3] Although we can note that ultrasound tests, the method used by most women for examining the child during pregnancy, is not completely reliable in determining child gender.

is no visible trend in the number of boys compared to girls (see Table A1 in Appendix). The percentage of boys is the same today (when abortions are free and gender detection methods more reliable), as it was in 1973 (when abortions were not free and child gender detection methods less reliable), indicating that sex-selective abortions are rare, if existent.

Larsson et al. (2002) also investigate the reasons for pregnancy termination among 518 Swedish women. In the questionnaire, they use open-ended questions, but none of the women stated child gender to be the reason for abortion. Instead, the most commonly stated reasons were financial concerns, bad timing or problems with the relationship.

4 Results

4.1 Graphical analysis

We start with a simple graphical analysis, plotting the distribution of parental leave days for mothers and fathers by child's gender (Figure 1). First, we can note the clear difference between mothers and fathers – while most mothers use a large amount of parental leave days, many fathers use only a few days. Second, and more importantly, the distribution of parental leave for parents with boys almost exactly mimics the distribution for girls (it is almost impossible to distinguish two different lines in the figure). Hence, from this rough graphical analysis, Swedish mothers and fathers seem totally unaffected by their child's gender when choosing the number of days on parental leave. Whether this conclusion holds for a more formal analysis is investigated below.

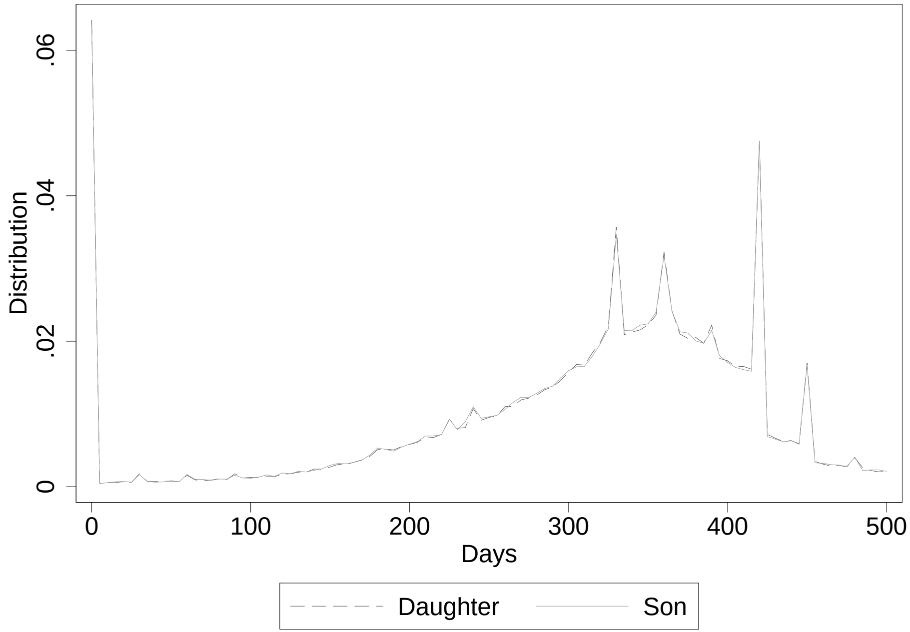

Figure 1 The distribution of parental leave days for mothers, by child gender

Note: For visibility, the graph is cut at 500 days; however, a small number of parents have used slightly more days since they got an additional child before the first child turned two years old.

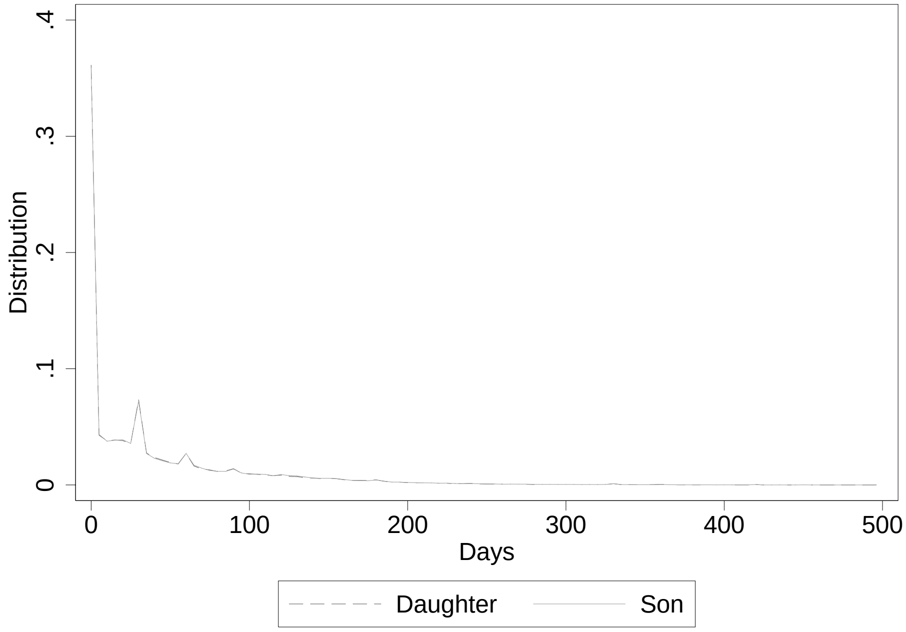

Figure 2 The distribution of parental leave days for fathers, by child gender

Note: For visibility, the graph is cut at 500 days; however, a small number of parents have used slightly more days since they got an additional child before the first child turned two years old.

4.2 Main results – incidence and duration

Table 2 shows OLS estimates of the effect of child gender on the probability of taking parental leave as well as the number of days on leave (given days>0). Clearly, there is no effect of child gender on the probability of taking parental leave, neither for mothers nor for fathers. However, among parents who do take some leave, child gender does seem to affect the length of the leave. On average, fathers are home slightly more than one half of a day longer with boys, while mothers are home an equally shorter period of time. (Results without covariates are shown in Table A2 in Appendix; clearly, the coefficients are almost unaffected, which is a further indication that child gender is exogenous.) Hence, children of each gender get approximately the same number of parental leave days, but fathers take a slightly larger fraction of this leave for sons.

Are these effects large? For mothers, definitely not. The daughter-baseline mean number of days on parental leave for mothers (during the child's first two years of life) is around 315 days; hence, the percentage change if the child instead is a son becomes -0.2 percent. For fathers, the change is larger: compared to the daughter baseline of 42 days, the increase for fathers having a son is 1.5 percent. This is well in line with the effect sizes found in for example Dahl and Moretti (2008) where daughters increase the probability that the child lives without a father by 3.1 percent and increases the probability that the mother has never been married by 1.4 percent. Hence, although the average effect of child gender on parental leave may be small in terms of how it affects child outcomes, it still shows gender biases among Swedish parents in about the same range as found in other industrialized countries.

Table 2 The probability of and days on parental leave for mothers and fathers

	Mother's pl>0	**Father's pl>0**	**Mother's pldays, given days>0**	**Father's pldays, given days>0**
Son	-0.000	0.001	-0.564*	0.639**
	(0.000)	(0.001)	(0.284)	(0.219)
Controls	Yes	Yes	Yes	Yes
R2	0.015	0.072	0.121	0.078
F	23.067	710.793	1451.275	710.742
N	356771	356771	353845	266759

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

4.3 Heterogeneous effects: different relative bargaining power

Is it the father, the mother, or both, that are responsible for this gender bias in parental leave? One way to investigate this issue is to show heterogeneous effects estimates for families with different relative bargaining power between the spouses. Assuming that the relative earnings and/or educational level of the mother is a relevant proxy for her bargaining power within the family, and assuming also that a higher bargaining power implies higher influence over the parental leave decision, this may shed some light on which spouse that is driving the results.

Table 3 and 4 show these heterogeneous effects estimates for mothers' and fathers' days on parental leave. Low (equal/high) maternal relative earnings mean maternal earnings below 33 percent (between 33 and 67 percent/over 67 percent) of family total earnings. Low (equal/high) maternal relative education means maternal education attainment that is lower than (equal to/higher than) paternal educational attainment. Note also that all standard control variables are included in these regressions. Most importantly, controls for both parents' absolute *levels* of education and earnings are included; this means that the variables for relative education/earnings should reflect bargaining power only and not the absolute levels of education and earnings.[4]

Perhaps suprisingly, the general pattern is that the gender bias is larger, the more bargaining power the mother has. For example, among non-traditional families where the mother has a higher level of education and/or earnings than the father, a first-born son increases fathers' parental leave by 1.3 to 1.6 days. This effect is more than two times larger than the average effect in Table 2 above. Among traditional families (with lower relative maternal earnings and/or educational levels), the effect of child gender is imprecisely estimated. This indicates that it may be mothers, rather than fathers, that are driving these gender biases. This is in clear contrast to previous studies who have traced the gender biases to fathers (see Dahl and Moretti, 2008). However, if we perform a Chow test to investigate if the "son"-coefficient is significantly different between the groups, the resulting F-statistics are in general small and we cannot reject the null hypothesis of no difference. There is one exception: the effect of having a first-born son on fathers' parental leave is significantly larger for families with higher maternal than paternal educational attainment than for families with lower maternal than paternal education (i.e. Table 4, panel b).

[4] In fact, heterogeneous effects estimates for different group of maternal and paternal educational *levels* show no statistically significant differences in gender bias at all. Focusing on earnings levels, there are no differences between different groups of paternal earnings, but among mothers, there is a concave pattern showing that the gender bias against sons diminish by income (high income mothers are less gender biased) but at a decreasing rate.

Table 3 Mothers' days on parental leave, given days>0: heterogeneous effects

	Low	Equal	High
		Panel a) Mothers' relative earnings	
Son	-0.091	-0.778*	-0.246
	(0.665)	(0.339)	(0.732)
Controls	Yes	Yes	Yes
R2	0.076	0.156	0.112
F	202.702	1199.631	185.965
N	83939	219420	50486
		Panel b) Mothers' relative education	
Son	-0.241	-0.498	-0.916
	(0.759)	(0.359)	(0.585)
Controls	Yes	Yes	Yes
R2	0.066	0.136	0.112
F	121.741	1089.216	326.763
N	56318	217037	80490

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

Table 4 Fathers' days on parental leave, given days>0: heterogeneous effects

	Low	Equal	High
		Panel a) Mothers' relative earnings	
Son	0.312	0.587*	1.638*
	0.535	0.244	0.675
Controls	Yes	Yes	Yes
R2	0.068	0.113	0.067
F	143.317	680.128	68.935
N	63538	174049	29172
		Panel b) Mothers' relative education	
Son	-0.171	0.588*	1.302**
	0.601	0.272	0.459
Controls	Yes	Yes	Yes
R2	0.058	0.098	0.050
F	73.957	598.017	109.267
N	40113	165867	60779

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

4.4 Effects over time

Given the relatively long panel (children born 1993-2005) we may investigate if the effect of child gender has changed over time. Gender roles are continuously changing and during this period, both the first and the second daddy-month were introduced, potentially leaving less room for personal gender preferences in the parental leave decision. The results when adding interaction terms between son and 1) being exposed to the first (but not second) daddy month and 2) being exposed to the second daddy month, are shown in Table 5 below (being born before both reforms is the baseline). There clearly seem to be a diminishing effect of child gender on mothers' parental leave, with the coefficient close to zero for families whose child was born after the second reform, but no change over time of the effect of child gender on fathers' leave.

Table 5 Effects over time

	Mother's pldays, given days>0	Father's pldays, given days>0
Son	-1.679*	0.646
	(0.725)	(0.711)
Son*daddymonth1	1.082	-0.063
	(0.823)	(0.766)
Son*daddymonth2	1.750*	0.069
	(0.883)	(0.804)
Controls	Yes	Yes
R2	0.121	0.078
F	1368.546	670.259
N	353845	266759

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

4.5 Longer run parental leave

Finally, we may also ask if these gender biases remain in the longer run. Table 6 shows estimates of the effect of the first-born child's gender on parental leave up to 8 years later. These estimates may still be regarded as causal, since the first-born child's gender is always exogenous. However, since most families get additional children, these estimates will also reflect the intermediate effect of the first-child's gender on fertility and the number of younger brothers and sisters.

As is clear from the table, the effect of having a first-born son is now imprecisely estimated (columns 1 and 4) and the point estimate is positive also for mothers, which might reflect the effect of a first-born son on fertility. The number of children in the family does, naturally, have a statistically significant effect on parental leave (columns 2 and 5) and when this (endogenous) variable is added as a control, the sign on the coefficient for a first-born son is again negative for mothers. If we instead of the son-dummy include a variable for the *share* of sons (columns 3 and 6) this variable is imprecisely estimated but the point estimates are in line with previous results, with small negative effects on mothers' leave and small positive effects on fathers' leave.

Table 6 Longer-run days on parental leave (up to child age 8), given days>0

	Mother's pldays, given days>0	Mother's pldays, given days>0	Mother's pldays, given days>0	Father's pldays, given days>0	Father's pldays, given days>0	Father's pldays, given days>0
Son	1.345	-1.169		1.002	0.728	
	(1.522)	(0.892)		(0.696)	(0.685)	
Children		299.326***	299.327***		29.276***	29.278***
		(0.984)	(0.984)		(0.675)	(0.675)
Shareofsons			-1.982			0.619
			(1.099)			(0.903)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.042	0.671	0.671	0.025	0.056	0.056
F	167.222	4484.969	4485.093	72.979	141.472	141.447
N	88099	88099	88099	77577	77577	77577

Notes: These regressions include families whose first child was born 1993-1995 (for younger children, long-term outcomes are not observed). Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

5 Preferences or constraints?

In economic modelling, any difference in behaviour is usually attributed to either preferences or constraints. In the context of parental gender bias towards the children, the constraints faced by parents of boys versus parents of girls could differ for two main reasons.

First, the monetary costs and benefits of boys and girls need not be equal. In many developing countries, boys imply larger economic returns to the parents since they for example provide old-age support. In industrialized countries, the costs of boys and girls could differ if boys for example are expected to require more expensive schooling.

Second, the production function of children could differ depending on gender for a number of reasons: fathers could be more important as role-models for boys than for girls, fathers could have a comparative advantage of raising boys, or boys could be more difficult to raise than girls (for example because they are more often sick), which may produce differences in parental behaviour towards children of different sexes also among parents without any gender preferences per se (see for example Dahl and Moretti, 2008). To empirically distinguish between the preference and the constraint explanation has been proven difficult since both give rise to almost identical predictions; see Lundberg (2005a). For example, if fathers prefer sons, the birth of a boy would increase marital stability simply because the father enjoys family life more in a boy-family. But the constraint explanation gives the same prediction – if a father feels more responsible for the upbringing of sons, then boy-births should increase marital stability.

Isolating the source of the gender bias – preferences, contraints, or both – lies beyond the scope of this paper. However, it is possible to shed some light on two of these alternative hypotheses, the monetary cost story and the sickness difference story.

5.1 Monetary costs

Different estimates of the cost of raising children in Sweden shows no or very small differences in the costs depending on gender (The Swedish Consumer Agency, 2001). The two main standards used by authorities to estimate child costs (in Swedish, "Riksnormen" and "Normalbeloppet") use the same cost estimates for boys and girls. A third estimate, developed by The Swedish Consumer Agency, does differentiate between girls and boys but find only small differences. Up to the age of ten, no differences are found. Between the age of 11 and 14, boys are on average found to cost 80 SEK (11.4 USD) more than girls each month, and boys aged between 15 and 18 years old are claimed to cost 230 SEK more each month (32.8 USD) (The Swedish Consumer Agency, 2001).

In relation to the total estimated cost of raising a child from birth to the age of 18, which lies in the range of around one million SEK (142 400 USD) in most estimates, these differences must be considered very small and unlikely to affect fathers' choice of whether or not to stay on parental leave. In addition, even if these minor differences would play a role, they would most likely lead us to underestimate the effect of child gender on parental leave. A natural guess is that the spouse with the highest income, usually the father, is at home less with boys since they are more expensive to raise, but the results goes in the opposite direction.

5.2 Morbidity differences

Theory suggests and empirical evidence shows that there are in fact small differences between newly born boys and girls (see Section 3.2). Table A3 in the Appendix shows some medical differences between newly born boys and girls in Sweden. Boys have larger head circumference than girls, and a larger percentage of boys are delivered by caesarean section. The mean time at hospital is also slightly longer, both for the mother and for the child, if the child is a boy compared to a girl. There are no differences in the percentage with a low APGAR score (a measure of the vitality of the child shortly after birth) between boys and girls, but a larger proportion of boys have some kind of malformation. Could these differences in birth complications and sickness be the single explanation of our findings?

One way to investigate this issue is to look at temporary parental leave usage (used for taking care of sick children, usually used for older children since it is only available for parents not on parental leave). To this end the data set Linda[5, 6] is used, a representative sample with register-based and longitudinal data on 3.35 percent randomly sampled Swedes and their family members. Now children of all parities, not only the first-borns, in ages 2-5 are included for natural reasons - temporary parental leave is mostly used for older children and focusing on first-borns without siblings would yield a selected sample of one-child families. This approach relies on the key assumption that siblings' gender should not affect the temporary parental leave decision by the parents. This is likely to be the case since parents on temporary parental leave typically are home only with the sick child, and not other siblings[7].

First, it is clear that the early differences in morbidity remain, at least during the child's first years of life. Comparing the total time spent on temporary parental leave for families with sons to families with daughters, we can see that boys remain sicker than girls (see Table A4 in Appendix)[8]. Second, I estimate the effect of child gender on fathers' days on temporary parental leave controlling for family total time on temporary parental leave for that specific child. In other words, I compare fathers' temporary parental leave for children with the same number of sick days. Note that the effect on mothers' temporary parental leave by construction is the exact opposite of the coefficient on father temporary parental leave (since we control for family total sick leave).

The results are shown in Table 7 below. Clearly, fathers' spend 0.085 more days on temporary parental leave for boys than for daughters. Since the daughter baseline mean sick leave is 2.59 days, this effect amounts to 3.3 percent, an even larger effect than in the main estimations above (and the effect is even larger if 2-year olds are excluded). Since we compare children with the same number of sick days, these results indicate that gender differences in morbidity are not enough to explain the difference in treatment of sons and daughters. The table also shows that these gender biases are of approximately the same magnitude regardless of the age of the child, with the exception of very young children (column 2).

[5] Longitudinal INdividual Data for Sweden

[6] For a more detailed description of the Linda database, see Edin and Fredriksson (2000).

[7] In contrast to ordinary parental leave, where a parent on leave typically takes care of all children in the family because child care access is partly restricted for children whose parents are on ordinary parental leave; see The Swedish National Agency for Education, 2001.

[8] In contrast to ordinary parental leave, which is given as a fixed amount of days for each child, the number of temporary parental leave days depends on the sickness of the child.

Table 7 Fathers' time on temporary parental leave

	All children 2-5 years old.	2-year olds.	3-year olds.	4-year olds.	5-year olds.
Son	0.085	0.029	0.127	0.085	0.106
	(0.021)***	(0.039)	(0.034)***	(0.029)***	(0.026)***
Total TPL, both parents	0.404	0.423	0.410	0.374	0.382
	(0.005)***	(0.009)***	(0.008)***	(0.010)***	(0.010)***
Constant	0.831	2.308	1.259	1.910	0.673
	(0.139)***	(0.284)***	(0.288)***	(0.326)***	(0.329)**
R-squared	0.44	0.45	0.44	0.41	0.42
N	235843	64405	60237	55796	55405

Notes: The outcome variable is the father's days on temporary parental leave during the years 1997-2003 for children aged 2-5 years with non-immigrant parents. The same standard control variables as in the main estimations above are included. Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

6 Concluding remarks

This paper investigates the presence of parental gender biases towards the children, as measured by the amount of parental leave used by mothers and fathers. A first-born boy increases fathers' time on parental leave by 0.6 days (1.5 percent) and decreases mothers' days on parental leave with a similar amount, leaving the total leave unchanged. From a policy perspective, these effects are small and unlikely to affect children's long-term outcomes. At the same time, the effects are in line with previous research on parental gender biases as measured by for example living arrangements and fertility, both in terms of sign and effect size. Hence, also in Sweden, a country with top ratings on gender equality, parental gender biases prevail.

Moreover, there are interesting differences between groups. Among non-traditional families, with high maternal relative earnings and/or educational levels, the effect of child gender is even larger. Assuming that maternal relative earnings/educational levels are reasonable proxies for maternal bargaining power, one interpretation of these results is that it is mothers, rather than fathers, that are responsible for the gender bias towards the children. This is in clear contrast to previous studies who mostly trace gender biases to fathers.

Finally, the present paper also investigates the effect of child gender on temporary parental leave, used for taking care of sick children, controlling for the total time sick for each child. By comparing children with the same number of sick days, we can rule out the competing hypothesis of gender differences in sickness as the sole explanation for any difference in parental behavior by child gender. Also these estimates show parental gender biases in the same direction and of slightly larger magnitude than for ordinary parental leave. Hence, gender differences in sickness are not enough to explain the differential treatment of sons and daughters by Swedish parents.

References

Almond, D. and L. Edlund (2006): Trivers-Willard at birth and one year, working paper 06-09, Institute for Social and Economic Research and Policy, Columbia, USA

Andersson, G., K. Hank, M. Rönsen and A. Vikat (2004): Gendering the family composition: sex preferences for children and childbearing behavior in the Nordic countries, Demography, vol. 43, no.2

Averett, S.L:, L.A. Gennetian, H.E. Peters (2005): Paternal child care and children's development, Journal of Population Economics, vol. 19

Batljan, I., S. Tillander, S. Örnhall Ljung and M. Sjöström (2004): Föräldrapenning, pappornas uttag av dagar, fakta och analys, Ministry of Health and Social Affairs (Socialdepartementet)

Bennedsen, M., K. Meisner Nielsen, F. Perez-Gonzalez and D. Wolfenzon (2007), Inside the family firm: the role of families in succession decisions and performance, Quarterly Journal of Economics, vol. 122, no 2

Berger, L.M, J. Hill and J. Waldfogel (2005): Maternity leave, early maternal employment and child health and development in the US, The Economic Journal, vol. 115

Bernal, R. (2008): The effect of maternal employment and child care on children's cognitive development, International Economic Review, vol. 49, no. 4

Blau, F.D and A.J Grossberg (1992): Maternal labor supply and children's cognitive development, Review of Economics and Statistics, vol 74, no 3

Cagnacci, A., A. Renzi, S. Arangino, C. Alessandrini and A. Volpe (2004): Influences of maternal weight on the secondary sex ratio of human offspring, Human Reproduction, vol. 19, no. 2

Catalano, R., T. Bruckner, J. Gould, B. Eskenazi and E. Anderson (2005a):, Sex ratios in California following the terrorist attacks of September 11, 2001, Human Reproduction, vol. 20, no. 5

Catalano, R., T. Bruckner, E. Anderson and J. Gould (2005b): Fetal death sex ratios: a test of the economic stress hypothesis, International journal of Epidemiology, no. 34

Chahnazarian, A (1988): Determinants of the sex ratio at birth: review of recent literature, Social Biology, Fall-Winter 1988

Chun, H. and J. Oh (2002): An instrumental variable estimate of the effect of fertility on the labour force participation of married women, Applied Economics Letters, vol. 9

Cox, Donald (2007): Biological basics and the economics of the family, Journal of Economic Perspectives, vol. 21, no. 2

Cunha, F. and J. Heckman (2007): The technology of skill formation, American Economic Review, vol. 97, no. 2

Dahl, G. B., and E. Moretti (2008), The demand for sons, Review of Economic Studies, vol. 75, issue 4

Desai, S, P.L. Chase-Lansdale and R.T. Michael (1989): Mother or market? Effects of maternal employment on the intellectual ability of 4-year old children, Demography, vol. 26

Doyle, O. , C.P. Harmon, J.J. Heckman and R.E Tremblay (2009): Investing in early human development: timing and economic efficiency, Economics and Human Biology, vol. 7, no. 1

Edin, P-A, and P. Fredriksson (2000): Linda – Longitudinal Individual Data for Sweden, working paper 2000:19, Department of Economics, Uppsala university

Larsson, M., G. Aneblom, V. Odlind and T. Tydén (2002): Reasons for pregnancy termination, contraceptive habits and contraceptive failure among Swedish women requesting an early pregnancy termination, Acta Obstetricia et Gynecologica Scandinavia, vol. 81, no. 1

Lundberg, S., S. McLanahan and E. Rose (2007): Child gender and father involvement in fragile families, Demography, vol. 44, no. 1

Lundberg, S. (2005a), Sons, daughters, and parental behaviour, Oxford Review of Economic Policy, vol. 21, no. 3

Lundberg, S. (2005b), The division of labor by new parents: does child gender matter?, Discussion Paper no. 1787, Institute for the Study of Labor, Bonn, Germany

Parcel, T.L and E. Menaghan (1994): Early parental work, family social capital and early childhood outcomes, American Journal of Sociology, vol. 99, no. 4

Ruhm, C. J. (2000): Parental leave and child health, Journal of Health Economics, no. 19

Ruhm, C. (2004): Parental employment and child cognitive development, Journal of Human Resources, vol. 39, no. 1

Tanaka, S. (2005): Parental leave and child health across OECD countries, The Economic Journal, vol. 115, no. 501

The National Board of Health and Welfare (2002): Fakta om mammor, förlossningar och nyfödda barn, report no. 2002:125:12

The National Board of Health and Welfare (2006): Aborter 2005, Statistik, Hälsa och sjukdomar 2006:2

The Swedish Consumer Agency (2001): Barn kostar. Räcker familjernas pengar? En jämförelse mellan olika normer., report no 2001:3

The Swedish Social Insurance Agency (2004a): Flexibel föräldrapenning – hur mammor och pappor använder föräldraförsäkringen och hur länge de är föräldralediga, report no 2004:14

The Swedish Social Insurance Agency (2004b): Föräldrapenningen – en översikt för åren 1974-2003, Statistikinformation Is-l:2004:4

The Swedish Social Insurance Agency (2005), Socialförsäkringstaket och föräldralön – ekonomi vid föräldraledighet, report no 2005:7

The Swedish National Agency for Education (2001): Plats utan oskäligt dröjsmål, report no. 2001:1944

Thomas, D. (1994), Like father, like son; like mother, like daughter: parental resources and child height, The Journal of Human Resources, vol. 29, no. 4, Special Issue: The Family and Intergenerational Relations

Trivers, R. L. and D. E. Willard (1973): Natural selection of parental ability to vary the sex-ratio of offspring, Science, no. 179

United Nations (1995): Human development report 1995: gender and human development, Oxford University Press

Wells, J.C.K. (2000): Natural selection and sex differences in morbidity and mortality in early life, Journal of Theoretical Biology, vol. 202, no.1

Appendix

A1 Additional tables

Table A1 Number and percentage newly born boys and girls in Sweden 1973-2005

	Boys		Girls	
Year	**N**	**Percent**	**N**	**Percent**
1973	56310	51,47	53092	48,53
1974	56673	51,50	53367	48,50
1975	53191	51,36	50368	48,64
1976	50542	51,39	47810	48,61
1977	49524	51,55	46548	48,45
1978	47819	51,33	45335	48,67
1979	49252	51,34	46685	48,66
1980	49732	51,41	47012	48,59
1981	47913	51,15	45765	48,85
1982	47398	51,47	44688	48,53
1983	47017	51,57	44149	48,43
1984	47809	51,46	45095	48,54
1985	50438	51,55	47409	48,45
1986	52069	51,39	49254	48,61
1987	53306	51,18	50844	48,82
1988	57493	51,58	53972	48,42
1989	59202	51,43	55910	48,57
1990	62876	51,28	59728	48,72
1991	63620	51,49	59935	48,51
1992	63166	51,45	59611	48,55
1993	59887	51,14	57215	48,86
1994	56737	51,08	54339	48,92
1995	52554	51,46	49577	48,54
1996	48369	51,06	46359	48,94
1997	45746	51,35	43339	48,65
1998	44319	51,63	41518	48,37
1999	44122	51,25	41967	48,75
2000	46226	51,56	43430	48,44
2001	46668	51,55	43854	48,45
2002	48364	51,33	45861	48,67
2003	50159	51,44	47352	48,56
2004	51739	51,51	48702	48,49
2005	51340	51,33	48677	48,67
Average		51,40		48,60

Source: Medical birth registry, The National Board of Health and Welfare. (The Medical birth registry covers virtually all births taken place in Sweden from early 1970s and onwards.)
Notes: This table includes all births (not only first births as I use in the estimations).

Table A2 The probability of and days on parental leave. Without covariates.

	Mother's pl>0	Father's pl>0	Mother's pldays	Father's pldays
Son	-0.000	0.002	-0.638*	0.700**
	(0.000)	(0.001)	(0.303)	(0.228)
R2	0.000	0.000	0.000	0.000
F	0.146	1.411	4.431	9.440
N	356771	356771	353845	266759

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Robust standard errors in parentheses.

Table A3 Medical differences between newly born boys and girls in Sweden, averages during 1973-2005

Head circum-ference (cm)		Low (0-6 p) APGAR score (%)		Malformation (%)		Mean hospital treatment time, mother (days)		Mean hospital treatment time, child (days)		Ceasarian section (%)	
Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
35,0	34,4	8,3	8,3	4,6	3,8	4,6	4,5	5,0	4,9	12,6	11,9

Source: Medical birth registry, The National Board of Health and Welfare.(The Medical birth registry covers virtually all births taken place in Sweden from early 1970s and onwards.)

Table A4 Parents mean number of days on temporary parental leave for boys and girls

	Boys		Girls		
	Mean no. of days	N	Mean no. of days	N	Significant difference
Days on TPL when child 2-3 years old	7.66	44820	7.20	42266	***
Days on TPL when child 3-4 years old	7.13	39124	6.84	37029	***
Days on TPL when child 4-5 years old	6.61	33797	6.33	31888	***
Days on TPL when child 5-6 years old	6.15	32741	5.83	30787	***

Notes: The outcome variable is the family total days on temporary parental leave during the years 1997-2003 for children with non-immigrant parents. Significance levels for difference in means: *** 1%, **5%, *10

A2 Details of the parental leave benefits over the years

Period	SGI days	% of income reimbursed	"Roof" of yearly income (SEK)	Max SEK/day, SGI days	Max SEK day if SGI=0	Flat rate days	SEK/dat, flat rate days
1990	360	90	222750	549	60	90	60
1991	360	90	241500	595	60	90	60
1992	360	90	252750	623	60	90	60
1993	360	90	258000	636	60	90	60
1994[a]	360	90	264000	651	64	90/0	60/0
1995[b]	360	80	267750	587	60	90	60
1996[c]	360	75	271500	558	60	90	60
1997	360	75	272250	559	60	90	60
1998	360	80	273000	598	60	90	60
1999	360	80	273000	598	60	90	60
2000	360	80	274500	602	60	90	60
2001	360	80	276750	607	60	90	60
2002[d]	390	80	284250	623	120	90	60
2003	390	80	289500	635	150	90	60
2004	390	80	294750	646	180	90	60
2005	390	80	295500	648	180	90	60
2006 (to June 30)	390	80	297750	653	180	90	60
2006 (from July 1)	390	80	397000	870	180	90	180
2007	390	80	398567	874	180	90	180
2008	390	80	397700	872	180	90	180
2009	390	80	415160	910	180	90	180

Notes: a) During the second half of 1994, the flat rate days were temporarily abolished for children >1 year old.
b) The first "daddy month" was introduced for children born after the 1st of january, 1995. During the 30 days set aside for each parent (the daddy month), the reimbursement level for the SGI days was still 90% of previous income.
c) During the 30 days set aside for each parent (the daddy month), the reimbursement level for the SGI days was still 85% of previous income.
d) The second "daddy month" was introduced for children born after the 1st of january, 2002.

Essay 4: The effect of own and spousal parental leave on earnings[1]

1 Introduction

The last decades have seen a convergence in the labor market behavior of males and females, where the male-to-female ratio of educational levels, participation rates, hours worked and hourly earnings have declined (Lundberg and Pollak, 2007; Lundberg, 2005). Despite this, females continue to take the lion's share of housework, child minding and parental leave (Evertsson and Nermo, 2007; Gershuny and Robinson, 1988; Halleröd, 2005; Lundberg and Pollak, 2007), and it is sometimes argued that this is one potential explanation for the remaining, unexplained earnings gap (Datta Gupta et al, 2008; Lundberg and Pollak, 2007). For example, being on parental leave for young children may reduce future earnings through a number of channels such as human capital losses during the absence period or signaling effects (Albrecht et al, 1999; Mincer, 1974; Mincer and Polachek, 1974; Mincer and Ofek, 1982; Stafford and Sundström, 1996). An additional mechanism, generally ignored in previous work, is the effect via future division of intra-household labor and child care. If parental leave today affects child care and household labor tomorrow, also spousal parental leave may be an important determinant of future earnings. For example, if a fathers' parental leave helps him acquire skills useful for taking care of children, this may affect future division of housework and child care within the family, and hence feed back onto maternal labor market behavior.

This paper investigates the effect of parental leave on earnings.[2] It fits into a broader literature on the effects of career interruptions on earnings.

[1] I wish to thank supervisors Peter Fredriksson and Per Johansson for excellent guidance. Helpful suggestions from Mikael Elinder, Jonas Lagerström, Håkan Selin, Björn Öckert and seminar participants at the Department of Economics, Uppsala university and participants at ELE conference on family economics in Lofoten are also acknowledged.

[2] The present study also serves to evaluate the Swedish daddy month reform. The main goals of the Swedish parental leave system are, as described in a government bill from 1993, gender equality, the child's right to both parents, child development and equal opportunity for both males and females to combine parenthood with a career (The Swedish Government, 1994). To my knowledge, there are no studies on how the daddy month affected parental labor market behavior.

However, the present paper departs from previous studies in several ways. First, it explicitly investigates the effect of not only own, but also spousal parental leave, an issue generally ignored in previous work. Second, it utilizes several sources of variation to identify effects. Besides cross sectional (CS) and fixed-effects (FE) models, it utilizes two policy reforms of the Swedish parental leave system that produced arguably exogenous variation in parental leave. The reforms reserved one and two months of leave for each spouse, which in practice decreased mothers' leave (the first reform) and increased fathers' leave (both reforms). Since the new rules applied to parents with children born after certain dates, the effect of reform exposure can be estimated using a difference in differences (DD) or triple differences (DDD) strategy. Finally, the register-based data set encompasses the entire Swedish population and is virtually free from missing-variables problems, attrition and self-report errors.

Previous studies have mostly found negative effects on earnings of absence in general and parental leave in particular (see for example Albrecht et al, 1999; Datta Gupta and Smith, 2002; Gangl and Ziefle, 2009; Görlich and De Grip, 2009; Mincer, 1974; Mincer and Polachek, 1974; Mincer and Ofek, 1982; Ruhm, 1998; Skyt Nielsen, 2009). In general, regression adjustment approaches are used for identification, sometimes with fixed effects to control for unobserved but time-invariant heterogeneity (Skyt Nielsen, 2009, is an exception using a reform of parental leave schemes among Danish publicly employed as exogenous variation). Regarding the effect of spousal parental leave, this issue is mostly ignored (one exception is Pylkkänen and Smith, 2003, who find that an increased parental leave period for fathers ("fathers' quota") reduces the job absence time of mothers, even when the days available for mothers are left unchanged). However, there are indications that early paternal involvement in childcare has effects on their involvement also later on. For example, Nepomnyaschy and Waldfogel (2007) find that fathers who take longer leave in connection to the birth of the child are more involved in child-caring activities 9 months later. On the other hand, Ekberg et al (2004) find no effects of ordinary parental leave on later care for sick children.

This paper shows that both own and spousal parental leave is potentially important for future earnings. Using the fixed effects model to control for unobserved but time-constant heterogeneity, the results show that each parent's own leave has a significant and negative effect on own future earnings. However, and more interesting, also spousal leave is important, but only for mothers. Each month the father stays on parental leave has a larger positive effect on maternal earnings than a similar reduction in the mother's own leave. Using the reforms as exogenous variation in parental leave yields imprecise estimates, despite the fact that both reforms strongly affected parental leave usage. However, the point estimates tentatively suggest larger effects than what was found using the fixed effects model.

2 The Swedish parental leave system and the reforms

The modern Swedish parental leave system was introduced in 1974, when both parents were given equal rights to use the system. It consists of several parts, the most important one being the governmentally paid cash benefit for parents staying home to care for their child. Most days (360 or 390, depending on child birth date) are reimbursed as a percentage of the previous wage, while a smaller amount of days (90) are reimbursed on a low flat rate. For individuals without the required previous labor market attachment, all days are replaced on a fixed (low) flat rate. The number of days on cash benefits as well as the reimbursement level has varied slightly over time; see Appendix for more details. There is great flexibility in the parental leave cash benefits; they can be used until the child turns eight years old and the parents can also choose to stay home part-time. The leave is also job protected. For more information on the Swedish parental leave system, see Berggren (2005), Duvander et. al. (2005) or The Swedish Social Insurance Agency (2002).

The overwhelming majority of parental leave is taken by mothers (Batljan et al, 2004). To increase the fathers' take up of parental leave benefits, two so called "daddy months" were introduced, the first in 1995 and the second in 2002. Before 1995, each parent were given half of the cash benefits days, but were free to transfer days to each other. But for those with children born from the 1st of January, 1995, 30 days of cash benefits are set aside for each parent and cannot be transferred. If those days are not used, they are simply lost. The 1st of January, 2002, an additional daddy month was introduced, making 60 days non-transferable. An important difference between the reforms is that in 1995, the total number of days was held constant, which meant that in practice mothers lost one month of parental leave. In 2002, the total number of days increased by one month so that mothers' maximum number of days was left unchanged.

It is important to note that the new rules apply according to the birth date of the child. There are also other changes in the parental leave system and in the social insurance system in general imposed from the 1st of January 1995 and the 1st of January 2002, but they generally apply equally to all individuals regardless of child birth dates. Hence, they affect both treatment (born after the turn of the year) and control (born after the turn of the year) groups equally. There are, however, some exceptions. The reimbursement rate was lowered from 90 to 80 percent in 1995. Although this affected all families equally in the long run, parents with children born before 1995 were given a respite and could keep their previous, higher replacement rate until the end of 1996. However, the 30 days set aside for each parent were excluded from this change and still replaced as 90 percent of previous wage. In 2002, the reimbursement rate for the flat rate days was doubled and this only applied to children born after 1st of January, 2002.

The daddy month legislation applies only to parents with shared custody of the child. Married parents are automatically given shared custody, while non-married parents must apply for shared custody. However, the overwhelming majority of families have shared custody. Within our sample (described below) 93 percent of all children had cohabiting parents at the time they turned one, and among cohabiting parents shared custody is very common. For example, 96 percent of all cohabiting parents of 1-5 year old children had shared custody in 1999 (Statistics Sweden, 2000). In the data, there is no information on custodial arrangements.

3 Identification

Theoretically, career interruptions and parental leave could affect an individual's own future earnings through three main channels. First is the effect via decreased market human capital (Mincer, 1974; Mincer and Ofek, 1982). This loss in market human capital may arise for different reasons such as a) forgone experience, b) skill depreciation during the leave, and c) effects ex ante via sorting into different types of jobs because of anticipated future career interruptions (Gronau, 1988). Second, career interruptions may work as a negative signal of work commitment (Albrecht et al, 1999; Datta Gupta and Smith, 2002). Third, there may be statistical discrimination against high absence groups (Gangl and Ziefle, 2009; Spence, 1973).

In addition, it is possible that not only the individuals' own but also *spousal* parental leave affects earnings. This possibility has generally been ignored in previous work. If we consider a standard model for intra-family division of labor, it implies that increasing returns to specialization, along with (possibly small) initial differences in (different types of) human capital endowments will induce females to at least partly specialize in home production and males in market work. This in turn lowers female annual earnings primarily via the direct effect on hours worked, but also via the effect on hourly earnings, as housework is assumed to lower hourly earnings through different channels (less effort left for work, less experience and human capital accumulation when working part-time or because of periods of job absence[3]) (Albrecht et al., 1999; Becker, 1991; Datta Gupta et al, 2008; Lundberg, 2005: Lundberg and Pollak, 2007, Mincer and Ofek, 1982; Stafford and Sundström, 1996). If the division of parental leave affects spousal relative human capital endowments, it could also affect earnings. For example, fathers on parental leave could acquire child care human capital if the parental leave implies a period of learning to take care of a child (this is especially likely if we focus on the first-born child) making him more likely

[3] Empirical support for this hypothesis is found in Hersch and Stratton (1994, 1997, 2000).

to take part of child care also in the future, which in turn could feed back to mothers' labor market behavior.

In the following, we focus on the effect of parental leave on mothers' earnings in a setting with panel data on families with their first child born in December or January around the reform cutoff or one year earlier.[4] Each family is observed twice, one year before birth and four years later. A flexible structural model for the effect of parental leave on mothers' earnings may be written

$$\ln E_{itcm} = \beta^0 + m^0 MPL_{itcm} + f^0 FPL_{itcm} + \alpha_c + \alpha_t + \alpha_m + \alpha_{mt} + \alpha_i + e_{itcm}$$

where the subscripts denotes family (i), time in terms of (approximate) child age (t=0 or t=4), cohort group (c=1 if the child is born around the reform cutoff) and month-of-birth (m=1 if born in January).

The dependent variable measures log earnings, MPL and FPL measures the mother's and the father's cumulative parental leave and the α:s denotes time (α_t), cohort (α_c), month-of-birth (α_m) and family (α_i) fixed effects. The interaction term α_{mt} allows the effect on earnings to vary between children born in December or January over time. This is potentially important, since we measure outcomes at the end of each calendar year. This means that children born in January are, by construction, on average one month younger when outcomes are measured than children born in December (remember that t denoted average child age; at t=4 children born in December are on average 4 years and 0.5 month old while children born in January are on average 3 years and 11.5 months old). This could imply that parents of January-born children are less likely to work or to work full-time and that those who do work are drawn slightly more from the upper end of the income distribution (the idea being that the reservation wage is higher, the younger the child is). This effect is also likely to vary over time – before birth (t=0) it is likely zero, while if we looked at t=1 it could be a sizeable effect and at t=4 it is probably smaller but perhaps not zero. Another example, which might produce systematic differences for parents of children born around the turn of the year, relates to the school starting age legislation. When children reach school starting age, there is a cutoff at the turn of the year, making children in the control group start school one year earlier than children in the treatment group which in turn could affect parent's labor market behavior[5]. However, this is probably a small concern since we measure outcomes for children below school starting age.

[4] Models for fathers' earnings may be written in an equal fashion but since the parameters may differ by gender the models need to be estimated separately for mothers and fathers.
[5] In Sweden, the mandatory school starting time is in August the calendar year when the child turns seven years old. One year earlier all children are offered to participate in a voluntary pre-school class during some hours each day. The pre-school classes are intended as a bridge

Since the family fixed effects are unobserved, we may rewrite $v_{itcm} = \alpha_i + e_{itcm}$ i.e. replace the error term and the family fixed effect with the composite error term v_{itcm}. For ease of exposition, control variables are omitted but can easily be added to the model. For simplicity we also disregard the fact that the number of parental leave days may enter nonlinearly; the intuition still holds for the more general case. Naturally, we would expect $|m|>|f|$, i.e that a mother's own parental leave have a larger effect on earnings than spousal parental leave. Previous research has generally ignored the spousal effects. However, here we have the explicit aim to estimate also the effect of spousal parental leave on own earnings.

First, if we only had cross-sectional data at t=4 the model would reduce to a standard cross-sectional (CS) model,

$$\ln E_{icm} = \beta^1 + m^1 MPL_{icm} + f^1 FPL_{icm} + \alpha_c + \alpha_m + \alpha_i + e_{icm} \qquad (1)$$

which is consistently estimated by ordinary least squares as long as $v_{icm} = \alpha_i + e_{icm}$ is uncorrelated with *MPL* and *FPL.* This assumption is unlikely to hold. For example, if parents who take more (less) parental leave also are less (more) career oriented and for that reason have lower (higher) earnings, this assumption is clearly violated. These differences in preferences for children versus market work may be difficult to proxy by including standard control variables and the resulting estimates will reflect selection rather than causal effects. In such case, the estimates will be biased downwards. Another possible story, potentially most applicable for fathers, is that fathers on leave – i.e. "responsible fathers" –are fathers with high earnings capacity. This interpretation is similar to the male marital wage premium found in earlier literature, where married men and/or fathers have higher earnings than non-married/non-fathers (Datta Gupta et. al, 2007; Gray, 1997). This story would lead to an upward biased estimate of the effect of parental leave on earnings among fathers.

Previous studies have used individual/family fixed effects to control for unobserved but time-invariant heterogeneity. If the endogenous variables – such as family preferences or "responsibility" – are constant over time, this approach yields unbiased estimates. Given our panel data, we can estimate a dummy-variable fixed effects (FE) model,

$$\ln E_{itcm} = \beta^2 + m^2 MPL_{itcm} + f^2 FPL_{itcm} + \alpha_c + \alpha_t + \alpha_m + \alpha_i + e_{itcm} \qquad (2)$$

---

between ordinary preschool and compulsory school (Swedish National Agency for School Improvement, 2007).

where we have assumed that $\alpha_{mt}=0$.[6] Note that *MPL* and *FPL* are always zero before birth so the main difference from model (1) above is that the dependent variable is measured as first differences. Now, the family unobserved effect can be controlled for so this model is consistently estimated by OLS under the weaker assumption $E[X_{itcm}*\Delta e_{itcm}]=0$, where X=*MPL*, *FPL*. In particular, the model allows for fixed family characteristics that are correlated with the dependent and independent variables.

However, to the extent that fertility (number, timing and spacing of children) is endogenous, also fixed-effects models may yield biased estimates (Browning, 1992; Lundberg, 2005). This could happen if, for example, fertility and/or parental leave respond to income shocks. If so, we need some kind of exogenous variation in parental leave to identify causal effects. This paper utilizes the daddy-month reforms as such plausibly exogenous variation and compares children born just around the reform cutoffs. If we continue to assume $\alpha_{mt}=0$ – i e. that there are no time-varying systematic differences between children born in December and January - we may restrict focus to children born around the reform cutoff only (and exclude families with children born the preceding year). Then a difference-in-differences (DD) model is given by

$$\ln E_{itm} = \beta^3 + rREFORM_{itm} + \alpha_t + \alpha_m + \alpha_i + e_{itm} \qquad (3)$$

where REFORM is an indicator variable for being exposed to the reform. Note that this variable is exactly the same as the interaction term between month-of-birth and time, α_{mt}, from above. This is why we need the $\alpha_{mt}=0$ assumption to hold in order for the REFORM coefficient to measure the effect of the reform (rather than the effect of differences between children born in December and January). If there are no such differences between children born in December and January, this model is consistently estimated by OLS as long as $E[REFORM_{itm*}e_{itm}]=0$. In particular, exposure to the reform should be exogenous and uncorrelated with for example income shocks. This specification identifies the intention to treat (ITT) effect – the effect of the reform on all families regardless of whether they comply or not – and as such, it mat be viewed as giving a lower bound on the "true" effect of a month increase/decrease in parental leave for fathers/mothers. In the absence of extra control variables, the reform coefficient equals the difference between different group means, see Table 1.

If there are normal-year systematic differences between children born in December and January ($\alpha_{tm}\neq0$), for example because children in the group exposed to the reform are slightly younger when earnings are measured, we

[6] Of course, we cannot distinguish between the different time-constant fixed effects, α_c,α_m and $\alpha_{i,}$ they are estimated simultaneously.

would need to include also families from a comparison year and estimate a difference-in-difference-in-differences (DDD) model,

$$\ln E_{itcm} = \beta^4 + r' REFORM_{itcm} + \alpha_c + \alpha_t + \alpha_m + \alpha_{mt} + \alpha_{ct} + \alpha_i + e_{itcm} \qquad (4)$$

where REFORM=α_{ctm} now is an indicator for children born in January during reform year at time t=4 (for completeness also the second "baseline" interaction effect α_{ct} is added to the model). In the absence of control variables, also this REFORM coefficient is given as a difference between group means; see Table 1 below.

Table 1 DD and DDD estimates

	Comparison group (child born one year before reform cutoff)		**Reform group** (child born around reform cutoff)	
Child's month of birth	**December**	**January**	**December**	**January**
lnE at t=0	a'	b'	a	b
lnE at t=4	c'	d'	c	d
Difference	c'-a'	d'-b'	c-a	d-b
DD estimate	(d'-b')-(c'-a')		(d-b)-(c-a)	
DDD estimate	[(d-b)-(c-a)]-[(d'-b')-(c'-a')]			

The models using the reforms as exogenous sources of variation (eq. 3-4) identifies the joint effect of MPL and FPL for the first reform, and the effect of FPL for the second reform. Remember that the second reform affected only fathers' parental leave while holding mothers' available parental leave days constant. In contrast, the first reform affected both parents' leave; given that mothers before the reform used virtually all parental leave, MPL was reduced by one month, while FPL was increased by a similar amount for the compliers.

Using the first reform, and without further assumptions about the parameters (m and f) we cannot identify whether the effect runs through own or spousal uptake of parental leave; we have only one instrument and two endogenous variables. But since we have two reforms, it is, in principle, possible to calculate instrumental variables estimates of the effect of each parent's parental leave (rather than the "reduced form" reform effects). However, such a strategy requires that there are no structural changes over time and since it is seven years between the first and second reform, this assumption may be questioned. We may also note that by using the reforms as exogenous variation and comparing families around the reform cutoffs our identification strategy isolates the direct and individual-level effect of parental leave on earnings. In particular, the estimated effect does not include long-term equilibrium effects, such as statistical discrimination, sorting into different types of job because of anticipated future job absence, or increased female investments in market human capital due to changed expectations of a future partner's share of housework.

For simplicity the discussion above did not include control variables. Given exogeneity of treatment status, control variables X are unnecessary; the inclusion of control variables may, however, increase precision and is also an informal way of testing exogeneity. Note, however, that the control variables are always measured prior to the child's birth and never in first differences even in the fixed-effects or DD/DDD models. (In the standard fixed-effects setting, non-variant control variables drop out; however, assuming that predetermined control variables can have different impact at different times/child ages allows us to include interactions between time and the pre-determined control variables.[7])

In the estimations, parental leave is measured only up to child age three (instead of four). The reason is that the outcome is annual earnings (as compared to wages or hourly earnings) and the prime purpose is to investigate the long-term effects of previous leave on future earnings (and not the obvious and immediate effect of parental leave today on earnings today). See Section A2 in Appendix for more details on the timing of variable collection.

As usual in earnings regressions, the problem of zeroes due to non-participation arises since we only observe earnings for individuals who participate in the labor market. Different processes may be at work on the extensive and intensive margin, and including observations with value zero and using a linear estimation model may induce specification bias due to nonlinearity. Focusing on individuals who do work necessarily implies conditioning on an endogenous variable which yields a selected sample of participants (Wooldridge, 2002). Throughout the paper OLS is used on log annual earnings in SEK+1 to include also non-participants but results on the participation decision as well as results for the participants only are shown in the Appendix.

4 Data

This section describes the data. It also describes how the reforms affected parental leave usage and discusses issues of exogeneity.

4.1 Data and estimation

The panel data set is based on register information (created by combining the LISA data base and the so called multigenerational registry, provided by

[7] Note that we never want to include control variables measured at t=4 since they may be affected by treatment and as such are part of the outcome.

Statistics Sweden, with data on parental leave provided by the Swedish Social Insurance Agency) encompassing the entire Swedish population. It contains high-quality, individual level information on all children and their family members, including information on annual earnings (from the tax registers), parental leave usage and standard covariates such as age, educational levels and marital status. There is in principle no attrition or missing variables problem.

The samples consist of native Swedish families[8] whose first child[9] was born one month before or after each reform cutoff or the preceding year. Families whose first birth was a multiple birth (approximately 3 percent) are excluded since the parental leave rules for these families are slightly different. This leaves us with 9007 families for the first reform sample and 8301 families for the second reform sample. In the main analysis, most variables are observed both one year prior to the child's birth (t=0) and when the child is on average four years old (t=4); see Section 3 above and Section A2 in Appendix for more details on the timing of data collection.

The dependent variable measures log annual earnings (in SEK + 1 to include zero-earners). The (possibly endogenous) independent variables of interest are the mother's and father's total parental leave up to child age three. These are measured in days in the descriptive section to give a precise picture of how the reforms affected parental leave, but for readability they are rescaled to months (by dividing by 30) in the regressions. These variables are used in models (1) and (2). The exogenous reform indicator, used in models (3) and (4), is 1 for children born in January 1995 (first reform sample) or January 2002 (second reform sample) and 0 for all other children. The models also include the other indicator variables mentioned in Section 3 (cohort, month-of-birth, time and their interactions). A number of control variables are also available, including parental age and educational levels, marital status and child gender.

Table 2 shows descriptive statistics for the samples. There are relatively small differences in terms of control variables both between comparison and reform periods and between children born in December and January. Most individuals have either a high school degree (around 60 percent) or a university degree (almost 30 percent). Fathers are older and have higher earnings than the mothers. A relatively small proportion (20 percent) is married and this is explained by the fact that marital status is measured one year prior to the child's birth.

Regarding parental leave, the reforms seem to have had a strong effect. The first reform decreased mothers' leave by around one month (27.8 days)

[8] For immigrants, there are around 20% missing observation due to lack of educational information. However, including immigrants in the estimations does not change the results.
[9] Only children who are *both* parents first-born child are included to avoid bias from previous children and their parental leave days.

and increased fathers' leave by almost 8 days. This is in clear contrast to the comparison period, where the number of parental leave days is quite similar for children born in December and January; slightly fewer days have been used for children born in January and that is probably because of the small difference in age. The second reform is associated with a decrease in mothers' leave by 10 days; however, in the comparison period mothers' days decreased by even more (14 days), which again suggests that this is due to the fact that children born in January are slightly younger than December-born children when parental leave is measured. Fathers' parental leave increased by 9 days after the second reform, while it remained virtually unchanged during the comparison period. These reform effects are slightly smaller than the ones estimated by Ekberg et al. (2005) and the reason is our focus on parental leave during the child's first three years of life.

Table 2 Descriptive statistics for the samples

	Comparison cohort		**Reform cohort**	
	Panel a) First reform sample			
	Dec93	**Jan94**	**Dec94**	**Jan95**
Mother's earnings	117.8	118.8	112.5	111.3
(thousands SEK)	(63.2)	(64.2)	(71.1)	(71.5)
Father's earnings	143.8	145.9	140.7	142.1
(thousands SEK)	(89.5)	(93.3)	(105.3)	(101.0)
Mother's PL (days)	460.4	457.8	467.1	439.3
	(160.6)	(154.6)	(159.7)	(152.5)
Father's PL (days)	50.1	47.5	40.4	47.9
	(69.4)	(69.7)	(71.2)	(62.0)
Mother's age	25.6	25.6	25.6	25.5
	(4.55)	(4.40)	(4.43)	(4.42)
Father's age	27.7	27.6	27.8	27.7
	(4.96)	(4.98)	(4.94)	(4.85)
Mother w. high school educ.	0.60	0.59	0.59	0.61
Father w. high school educ.	0.58	0.60	0.60	0.58
Mother w. university educ.	0.29	0.29	0.29	0.28
Father w. university educ.	0.27	0.26	0.27	0.28
Married	0.18	0.17	0.18	0.18
Son	0.51	0.51	0.53	0.51
N	2135	2520	2115	2237

	Panel b) Second reform sample			
	Dec00	**Jan01**	**Dec01**	**Jan02**
Mother's earnings	155.3	155.8	170.0	170.4
(thousands SEK)	(95.1)	(99.4)	(104.5)	(105.7)
Father's earnings	205.4	209.3	226.9	222.1
(thousands SEK)	(128.4)	(134.4)	(176.7)	(168.1)
Mother's PL (days)	408.1	394.4	405.3	395.2
	(142.5)	(142.6)	(146.9)	(138.8)
Father's PL (days)	56.6	57.3	62.5	71.6
	(69.2)	(68.5)	(67.9)	(69.7)
Mother's age	26.8	26.9	27.3	27.0
	(4.50)	(4.58)	(4.71)	(4.55)
Father's age	28.9	28.9	29.2	28.8
	(5.01)	(4.92)	(5.04)	(4.97)
Mother w. high school educ.	0.50	0.48	0.54	0.56
Father w. high school educ.	0.54	0.53	0.66	0.64
Mother w. university educ.	0.38	0.40	0.38	0.36
Father w. university educ.	0.34	0.35	0.24	0.26
Married	0.21	0.20	0.20	0.18
Son	0.53	0.52	0.53	0.51
N	1848	2174	1944	2335

Notes: All variables except the parental leave variables and child gender are measured one year prior to the child's birth. Earnings are measured in thousands SEK, including zeroes. Standard errors in parenthesis.

4.2 How the reforms affected parental leave use

As a start, it is illuminating to look at how the reforms affected parental leave use from different angles. Figure 1 starts by plotting the mean number of parental leave days (measured at the end of the calendar years three years after the birth-turn of the year) for different child birth month cohorts (December- or January-born children from different years). This shows the development of parental leave over time. Clearly, fathers' parental leave increased at both reform cutoffs, while mothers' parental leave decreased only at the first reform cutoff. However, mothers with children born in January seem to always have used slightly fewer parental leave days, most likely because their children are on average one month younger when outcomes are measured. This small difference in child age does not seem to affect fathers' parental leave during non-reform years.

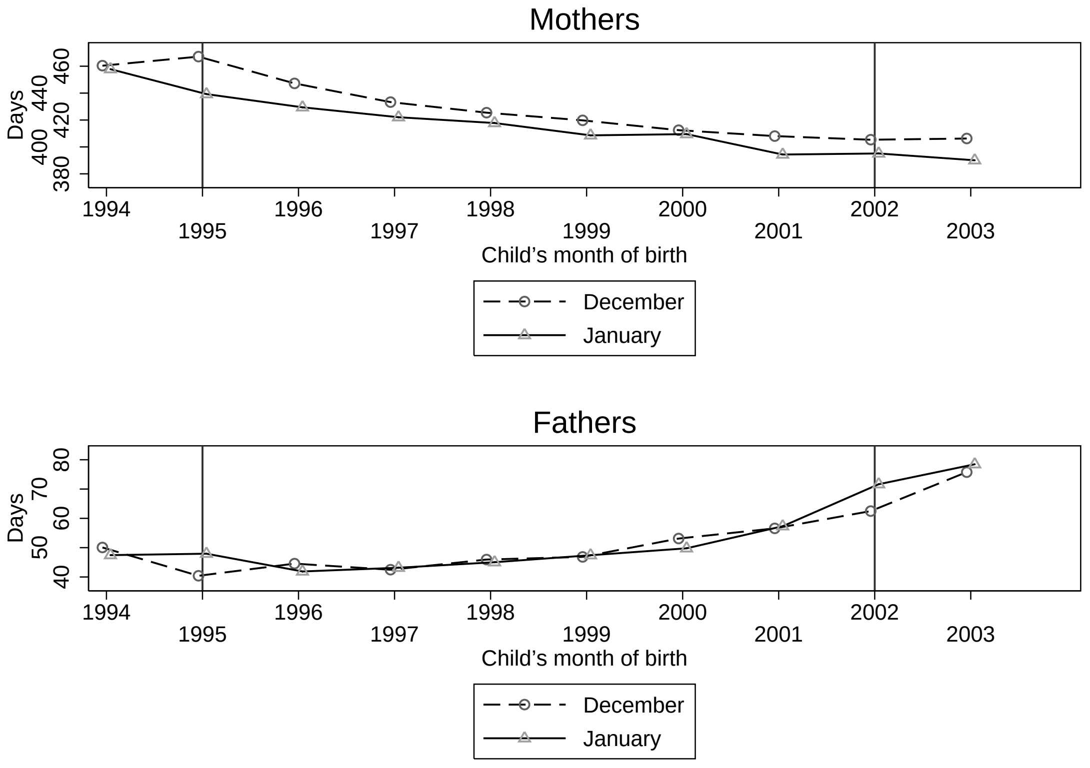

Figure 1 Mean parental leave days for different child mob-cohorts

Second, Table 3 shows the results when parental leave days are regressed onto reform exposure status with and without control variables (i.e. the DDD model (4) above but with mothers' and fathers' days on parental leave instead of earnings as dependent variable). Clearly, the reforms effectively increased fathers' leave by around 9-10 days each, and the first reform decreased mothers' leave by almost 26 days. The reform coefficients do not change much when control variables are added, which indicates that the reforms were exogenous to the parents. However, this issue is more deeply investigated in the Section 4.3 below.

It is also interesting to investigate if there are heterogeneous responses to the reform, i.e. to examine the compliers. Tables A1 and A2 in Appendix show the reform effects for subgroups with different levels of maternal and paternal education. Although the patterns are not so clear it does seem like both reforms had relatively smaller effects on fathers' leave among families with a low maternal level of education.

Table 3 The effect of the reforms on parental leave use

	Mothers' days	**Mothers' days**	**Fathers' days**	**Fathers' days**
		Panel a) First reform sample		
REFORM	-25.304**	-25.792**	10.144*	10.025*
	9.383	9.352	4.075	4.046
Controls	No	Yes	No	Yes
R2	0.905	0.905	0.595	0.602
F	9537.692	3222.543	537.277	190.664
N	18014	18014	18014	18014
		Panel b) Second reform sample		
REFORM	3.522	1.537	8.444*	9.019*
	8.910	8.817	4.289	4.213
Controls	No	Yes	No	Yes
R2	0.899	0.901	0.648	0.660
F	8188.341	2822.574	859.009	305.184
N	16602	16602	16602	16602

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Next, we take a closer look at the behavior around the reform cutoffs. Figure 2 shows the timing of parental leave for mothers and fathers in the reform cohorts (January 1995 versus December 1994 and January 2002 versus December 2001). More specifically, it shows the number of parental leave days each month during the child's first 6 years of life. Note that parental leave days from younger siblings show up in this figure since the parents get additional leave entitlements for each child.

Clearly, most days are used before the child turns two years old. For mothers, there are no clear seasonal patterns and no differences between the January (solid) and December (dashed) group except that the graph for

January-mothers in the first reform sample lies slightly below the graph for December-mothers, a natural result of the reform. For fathers, we may note several interesting features. First, the graph for the January group mostly lies above that for the December group, which indicates that the January group indeed used more parental leave. Second, there are clear seasonal trends – fathers seem to use more parental leave during holidays, primarily during the summers but also in connection with Christmas and New Year. That is also the most likely explanation for the small difference in timing between January and December groups – fathers in the January group are on parental leave slightly earlier and this may be because of the timing of holiday breaks. Apart from that, the differences between December and January groups are small.

Finally, Figure 3 shows the distribution of the amount of parental leave for January (solid) and December (dashed) group, respectively. (In this figure, parental leave is measured up to child age three – the variation that is used in the main analysis - but looking at longer run parental leave does not change the overall picture). For the first reform sample, the distribution of fathers' days is clearly shifted to the right as a result of the reform, with a new peak at around 30 days. The distribution of mothers' days is likewise shifted to the left (the peaks for mothers are located at or slightly below the maximum available days on benefits, with and without the flat rate days). In this picture, the second reform does not seem to have affected mothers' distribution of leave, but fathers' leave was again shifted to the right with a new peak at 60 days.

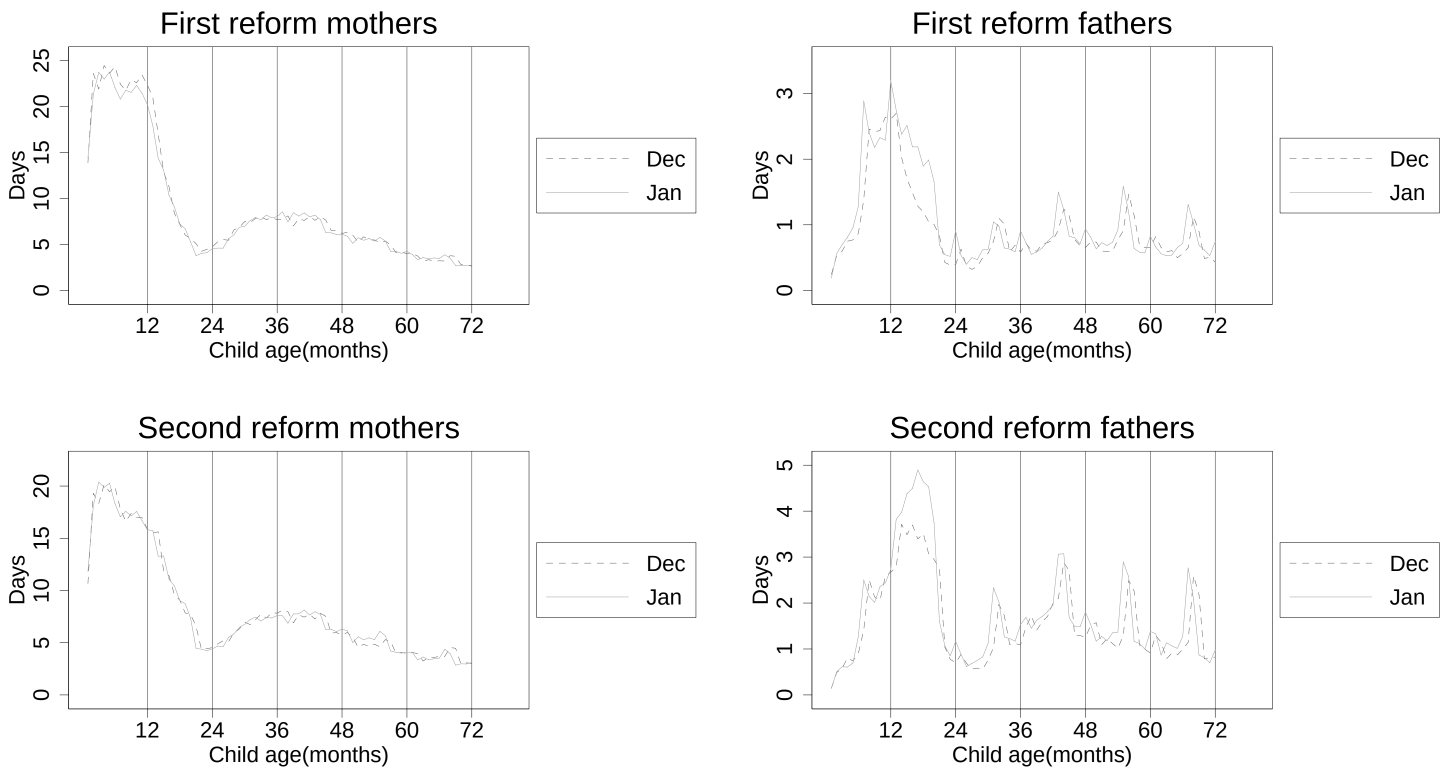

Figure 2 The timing of parental leave for reform cohorts by child month-of birth (December/January)

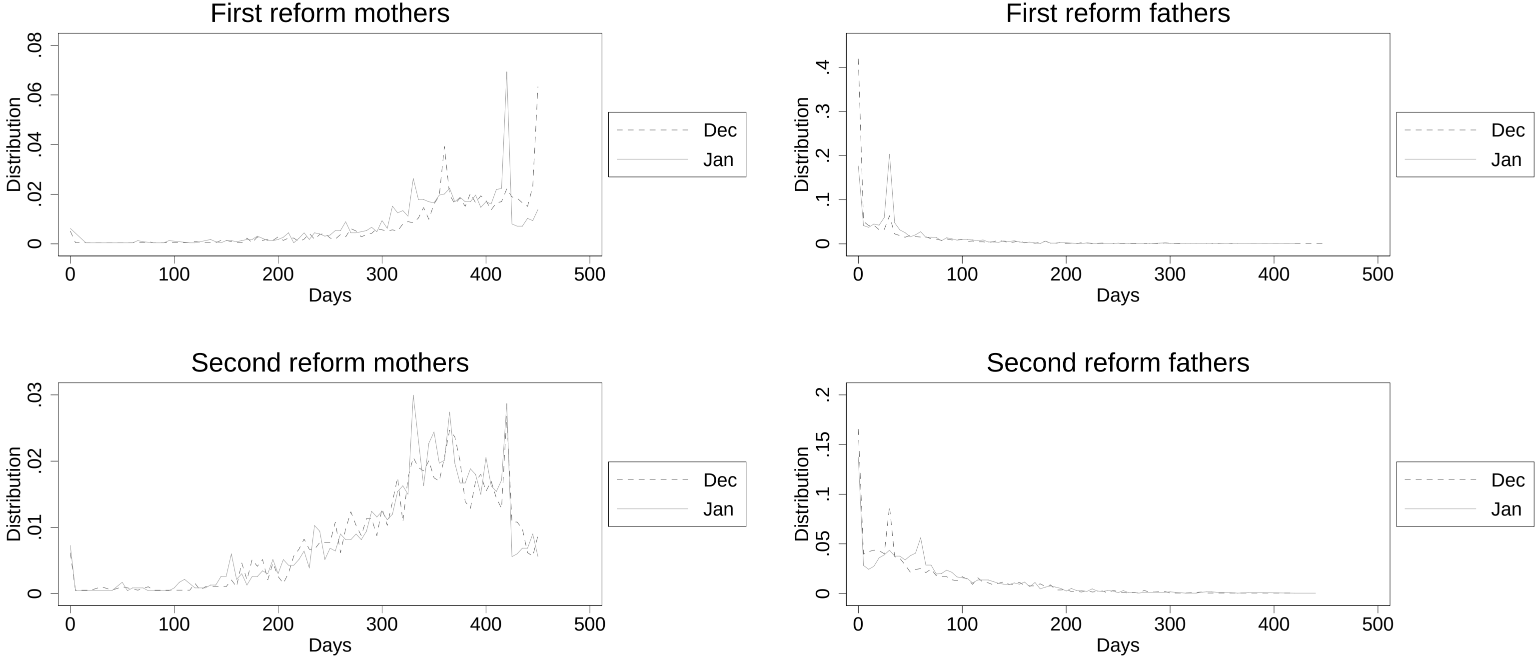

Figure 3 The distribution of parental leave days for reform cohorts by child month-of birth (December/January)

Note: for visibility, the graph is cut at the one-child maximum of 450 days; however, a smaller amount of parents have used slightly more days than this since they had another child before the first child turned three.

4.3 Exogeneity of reform exposure

The parental leave reforms in 1995 and 2002 are used as exogenous sources of variation in order to estimate the causal effect of parental leave on earnings. This identification strategy requires that a) no other change, affecting treatment and control groups differently, occurs at the same point in time as the reforms, and b) there is no endogenous sorting at the reform thresholds.

Regarding (a), are the reforms the single changes affecting January and December groups differently? Again, there were other changes in the social security system passed the 1st of January in 1995 and 2002, but they generally affected both groups equally. Only the daddy-month introduction along with some smaller changes in the reimbursement rate for the transferable days (*not* the daddy-month) was tied to the birth date of the child.

Regarding (b), is there any endogenous sorting at the reform thresholds? We start by investigating static sorting, although it is worth noting that fixed individual characteristics are allowed to be correlated with the probability of reform exposure (the individual fixed effects are differenced out; see Section 3 above). However, if there are static sorting it is also possible that there are sorting in terms of time-varying variables as well.

The first reform gave incentives for parents to induce an earlier birth, both to avoid the daddy month restriction and because of the slightly higher replacement rate for children born before 1995. The second reform reversely gave incentives to postpone birth since the parental leave rules were strictly better for children born after the reform. These incentives may have caused informed parents to fine-tune delivery. Are there such indications?

The first reform was difficult to anticipate at the time of conception. Although the daddy-month debate had been going on for years, it was unclear whether, when and how it should be implemented. As late as the 26th of April, 1994, three parties from the governing coalition threatened to vote against any such proposal (Karlsson, 1994a) and the reform proposition was not passed until 30th of May, 1994 (Karlsson, 1994b) when the turn of the year babies 1994/1995 were already conceived. Even so, parents could of course plan an earlier birth just in case. In addition, although the exact natural birth date is a random process it is in principle possible to induce an earlier birth by medical means, for example by using a caesarian section. The second reform had been known long in advance (TT, 2001) and informed parents may have chosen to postpone childbearing.

Since there may be room for sorting around the reform cutoffs, we investigate this issue a little deeper. First, Figure 4 below plots the number of first births in December and January over time. There clearly seem to be large variations over time, and possibly some tendencies of sorting in the anticipated direction – the difference in births between January and

December are relatively small in 1994/1995 and slightly larger in 2001/2002. However, such tendencies exist also at other points in time. In 1999/2000, for example, the difference is even smaller than in 1994/1995.

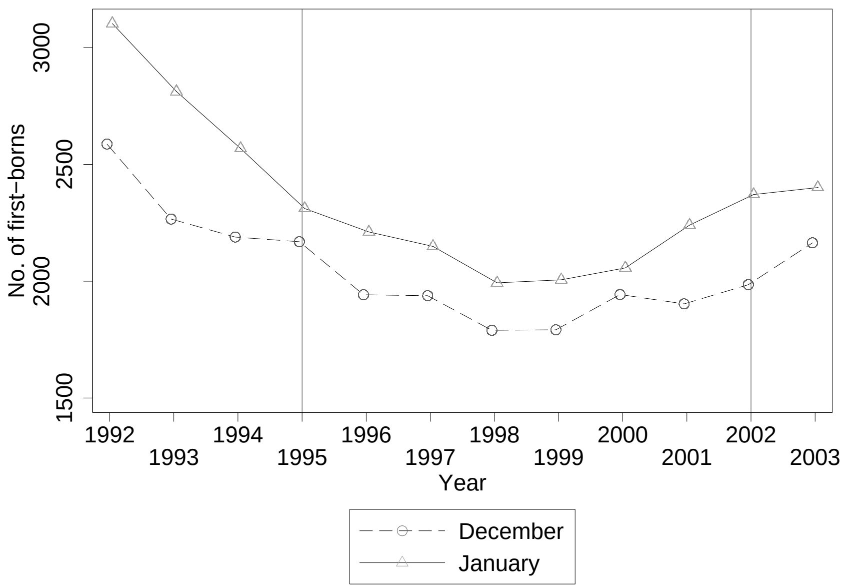

Figure 4 Number of first births in December and January over time

Next, we investigate whether observables can explain treatment status. This may show if there are indications of endogenous sorting at the reform cutoffs or if the pattern in Figure 3 above is merely the result of random variation. (Of course, there could be endogenous sorting that does not show up in terms of observables, but that is impossible to investigate). Table 4 shows regression results when an indicator variable for being exposed to the reform is regressed onto some arguably exogenous covariates (i.e. model (4) above but where the outcome variable is REFORM status and this is regressed onto all other fixed effects and the control variables).

Clearly, there are no statistically significant differences in parental characteristics[10] between children born in January and December and all point estimates are small in magnitude[11]. However, even if each single coefficient is statistically non-significant, they could have explanatory power together. In fact, F-tests between these models and similar models without

[10] See also Ekberg et. al. (2005) who compare the number of births each *day* around the turn of the year 1994/1995 and other years and find no systematic pattern. In addition, they compare parental age distributions for children born two *weeks* before and after the reform and find no evidence of differences in parental characteristics.

[11] All variables except the child gender variable are measured prior to the birth of the child.

control variables (only the fixed effects for cohort, time, month-of-birth and their pairwise interactions are included) returns test statistics of 2.77 (first reform sample) and 2.58 (second reform sample) which is statistically significant and rejects the null hypothesis that the added control variables have no explanatory power. So, there may be some static sorting in terms of observable characteristics. This suggests that there could also be sorting in terms of unobservables. However, as noted above, static sorting is in itself not problematic (since we have panel data and can estimate the family fixed effects).

Next, we investigate the more important issue, if there seems to be time-variant sorting. In particular, we do not want reform exposure to be correlated with income shocks. Instead, January and December groups should follow the same wage growth paths over time. Table 5 investigates this issue by regressing the probability of reform exposure (being born in January around the reform cutoff) on the fixed effects, the control variables and different earnings lags (maternal and paternal earnings two and three years before the birth of the child). This is necessarily done on a slightly smaller sample since these earnings lags are not available for all individuals. At most, we lose 74 individuals from the first reform sample and 75 individuals from the second reform sample. Clearly, none of the earnings lags are statistically significant and they are also small in magnitude. Hence, the groups exposed to the reforms seem to follow the same earnings pattern over time as the comparison groups.

Table 4 The effect of exogenous characteristics on prob(reform exposure)

	First reform	Second reform
Mother's lnE	(-0.002	0.002
	(0.001)	(0.002)
Father's lnE	0.001	-0.001
	(0.001)	(0.001)
Father's age	-0.000	-0.000
	(0.001)	(0.001)
Mother's age	-0.000	-0.001
	(0.001)	(0.001)
Father w. high school educ.	-0.012	0.003
	(0.011)	(0.013)
Mother w. high school educ.	0.011	0.001
	(0.013)	(0.015)
Father w. university educ.	-0.000	0.015
	(0.014)	(0.015)
Mother w. university educ.	0.005	-0.012
	(0.015)	(0.016)
Married	0.002	-0.001
	(0.010)	(0.010)
Son	-0.005	-0.003
	(0.007)	(0.008)
R2	0.857	0.872
F	1031.9	1329.9
N	18014	16602

Notes: All variables (except child gender) are measured one year before the birth of the child. Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table 5 The effect of income lags on prob(reform exposure)

	Prob (reform exposure)	Prob (reform exposure)	Prob (reform exposure)	Prob (reform exposure)
		Panel a) First reform sample		
Mother's lnE, lag2	-0.001			
	(0.002)			
Mother's lnE, lag3		-0.000		
		(0.002)		
Father's lnE, lag2			0.002	
			(0.001)	
Father's lnE, lag3				0.001
				(0.001)
Controls	Yes	Yes	Yes	Yes
R2	0.857	0.857	0.857	0.857
F	1110.1	1089.9	1114.3	1111.0
N	17970	17866	17998	17970

	Prob (reform exposure)	**Prob (reform exposure)**	**Prob (reform exposure)**	**Prob (reform exposure)**
		Panel b). Second reform sample		
Mother's lnE, lag2	0.002			
	(0.001)			
Mother's lnE, lag3		0.001		
		(0.001)		
Father's lnE, lag2			-0.000	
			(0.001)	
Father's lnE, lag3				-0.000
				(0.001)
Controls	Yes	Yes	Yes	Yes
R2	0.872	0.872	0.872	0.872
F	1429.7	1403.4	1434.8	1427.8
N	16522	16452	16565	16533

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

4.4 Preview of results – simple cross-tabulations

Without control variables, the REFORM-coefficient in the difference-in-differences (DD) and triple differences (DDD) models can be calculated as simple differences between group means. Tables 6 and 7 below shows these estimates for mothers and fathers; both estimates are also shown for different placebo years and the DDD-estimates are calculated using different comparison years. For ease of exposition, standard errors are omitted but as will be clear from Section 5.1 the standard errors are indeed huge and none of the differences below are statistically significant.

The first reform increased mothers' subsequent earnings by 9 percent using the DD approach and by 10-15 percent using the DDD approach with different comparison years. Hence, it is a sizeable positive effect of the first reform on mothers' earnings, and the point estimate also seems robust to different comparison years. In addition, the DD- and DDD-estimates from different placebo years are all much smaller and mostly of the reverse sign, which further indicates that the reform indeed had an effect on maternal subsequent earnings. However, turning to the second reform, the results are less robust. The coefficients from DD and DDD-models vary in both sign and size (from -5 percent using the DD model to between 1 and 11 percent using DDD-models) and the result are not very different from estimates in different pre-reform placebo years.

This could indicate that it is mothers' own leave (which was affected by the first but not the second reform) that is important. (Another possible story is that there could be differences in parental leave timing between the reforms. Potentially the first reform induced fathers to take more "non-holiday" parental leave, since otherwise the total expected leave was reduced, while the second reform was less strict in the sense that the families were given an additional month of leave, implying that fathers could more freely choose the timing of the parental leave. If so, and if "holiday"-parental leave is less helpful for maternal labor market behavior, this could explain the difference in effects between the first and second reform.)

Regarding the fathers, both reforms seem to have had a negative effect on subsequent earnings. The first reform's estimates range from -18 to -34 percent, indeed huge effects but suprisingly robust to the choice of comparison year and also more negative than any of the pre-reform placebo estimates. The second reform's estimates are much smaller, -5 to 5 percent, and also quite similar to the pre-reform placebo estimates.

Table 6 Cross tabulations with DD and DDD estimates, mothers

	Comparison cohort 3		**Comparison cohort 2**		**Comparison cohort 1**		**Reform cohort**	
				Panel a) First reform sample				
	Dec 91	**Jan 92**	**Dec 92**	**Jan 93**	**Dec 93**	**Jan 94**	**Dec 94**	**Jan 95**
LnE at t=0	11,21	11,20	11,11	11,05	10,96	11,00	10,65	10,50
LnE at t=4	9,29	9,24	9,39	9,32	9,49	9,48	9,64	9,57
Diff	-1,93	-1,96	-1,72	-1,73	-1,47	-1,53	-1,02	-0,93
DD estimate		-0,03		-0,01		-0,06		**0,09**
DDD estimate1				0,02		-0,05		**0,15**
DDD estimate2						-0,03		**0,10**
DDD estimate3								**0,12**
				Panel b) Second reform sample				
	Dec 98	**Jan 99**	**Dec 99**	**Jan 00**	**Dec 00**	**Jan 01**	**Dec 01**	**Jan 02**
LnE at t=0	10,65	10,65	10,93	10,95	11,08	11,05	11,25	11,31
LnE at t=4	10,16	10,00	10,22	10,18	10,18	10,01	10,06	10,07
Diff	-0,49	-0,65	-0,71	-0,77	-0,91	-1,05	-1,19	-1,24
DD estimate		-0,16		-0,06		-0,14		**-0,05**
DDD estimate1				0,10		-0,08		**0,09**
DDD estimate2						0,02		**0,01**
DDD estimate3								**0,11**

Table 7 Cross tabulations with DD and DDD estimates, fathers

	Comparison cohort 3		**Comparison cohort 2**		**Comparison cohort 1**		**Reform cohort**	
				Panel a) First reform sample				
	Dec 91	**Jan 92**	**Dec 92**	**Jan 93**	**Dec 93**	**Jan 94**	**Dec 94**	**Jan 95**
LnE at t=0	11,21	11,24	11,14	11,15	10,87	10,82	10,39	10,55
LnE at t=4	10,98	11,10	11,04	11,10	11,18	11,05	11,26	11,18
Diff	-0,24	-0,14	-0,10	-0,04	0,30	0,24	0,88	0,63
DD estimate		0,09		0,06		-0,07		**-0,24**
DDD estimate1				-0,04		-0,12		**-0,18**
DDD estimate2						-0,16		**-0,30**
DDD estimate3								**-0,34**
				Panel b) Second reform sample				
	Dec 98	**Jan 99**	**Dec 99**	**Jan 00**	**Dec 00**	**Jan 01**	**Dec 01**	**Jan 02**
LnE at t=0	10,93	10,91	11,05	11,16	11,29	11,34	11,51	11,38
LnE at t=4	11,58	11,48	11,46	11,59	11,50	11,58	11,68	11,53
Diff	0,64	0,57	0,42	0,43	0,21	0,24	0,17	0,15
DD estimate		-0,07		0,01		0,03		**-0,02**
DDD estimate1				0,09		0,01		**-0,05**
DDD estimate2						0,10		**-0,04**
DDD estimate3								**0,05**

5 Results

5.1 Main results

Table 8 and show estimation results for mothers and fathers for the first and second reform sample separately and using the different models (cross-section, fixed effects, DD and DDD).

There are several things to note. First, there are clear differences between the cross-sectional model and the fixed-effects model, which suggest selection of families into different levels of parental leave usage. Second, using the fixed-effects model, own parental leave do seem to reduce subsequent earnings – each month of own parental leave lowers mothers' earnings by 4.5 percent (in the first reform sample) and fathers' earnings by around 7.5 percent. The magnitude of these effects is far larger than previous studies – for example, Albrecht et al (1999) found wage reductions of 0.1-0.5 percent for each month of parental leave. This can be explained by the fact that here, annual earnings are used which reflect both wages and hours worked, while most previous studies have focused on wages. In addition, our focus is on the relatively short run effect on earnings four years later, when some parents could still be on parental leave (and parental leave up to child age 3 may be correlated with later parental leave). In addition, the longer-run effects are usually found to be smaller due to rebound effects and catching-up of human capital.

The differences in effects between males and females could be due to nonlinearities, if the first months of leave are more important for earnings than later parental leave. It could also be a signaling effect. As suggested by Albrecht et al (1999), parental leave could have a stronger signaling value for males since so few fathers stay on parental leave compared to virtually all mothers.

Third, and more interesting, spousal parental leave has no effect on father's earnings but do seem important for mother's labor market behavior. Each additional month that the father stays on parental leave increases mothers' earnings by 6.7 percent in the first reform sample (the effect in the second reform sample is not statistically significant). This is a large effect, even larger than the effect of a mother's *own* parental leave. This indicates that paternal (lack of) involvement in parental leave and child care may in fact be one important explanation for the male-to-female earnings gap. Another story could be a "reverse signaling" story – while most mothers take all available parental leave, a shorter period of leave could work as a positive signal of work-commitment.

These causal interpretations rest on the assumption of no time-variant unobserved heterogeneity, and in particular that fertility and parental leave is not endogenous. For example, if parents who experience an income shock becomes more (less) likely to have children and/or stay on parental leave, this assumption is clearly violated. Using the reforms as exogenous variation in parental leave do, unfortunately, yield very imprecise estimates that are not statistically different from zero. We can note, however, that this is not because of a weak effect on parental leave use. As we saw in Section 4.2, the reform effectively changed the parents' time on parental leave. Instead, it could be that the normal-year variation in earnings depending on child birth dates is too large to enable precise estimation.

However, we may still make some comparisons of the point estimates across models. The tables also report the predicted reform effect for the CS/FE-models, which is a calculation of the predicted effect of the reform if the assumptions underlying the CS or FE models are fulfilled. This effect is calculated as the mean change in mothers' and fathers' time on parental leave as induced by the reforms (see the reform-coefficient from Table 2 above, columns 2 and 4), multiplied by the coefficient on each month of leave as estimated by the CS/FE models.[12]

For example, if the fixed-effects results are true, we would expect the first reform to increase maternal earnings by 6.1 percent; both because of the decrease in own leave and because of the increase in spousal leave. This

[12] The standard error of this estimate is calculated assuming that the underlying variables are independent random variables.

effect is well within the 95 percent confidence interval of both models using the reform as exogenous variation. The most flexible model, DDD, tentatively suggests even larger effects – the point estimate is 14.9, albeit very imprecisely estimated. The same pattern is found also for the second reform sample and among fathers – model (4) always returns larger point estimates than model (2). This tentatively suggests that the "true" effect is in the same range or larger than suggested by the fixed-effects specification.

Finally, we can note that these estimates are quite similar to the estimates without control variables (see the cross-tabulations above), which further indicates that the reforms are indeed exogenous.

Table 8 The effect of parental leave on mothers' earnings at child age 4.

	CS	**FE**	**DD**	**DDD**
		Panel a) First reform sample		
Mother's PL	-0.011	-0.045***		
	(0.009)	(0.013)		
Father's PL	0.021	0.067*		
	(0.019)	(0.029)		
REFORM	[0.017]	[0.061]	0.088	0.149
	[0.011]	[0.023]	(0.176)	(0.244)
Controls	Yes	Yes	Yes	Yes
R2	0.059	0.656	0.667	0.655
F	40.717	45.833	17.038	41.939
N	9007	18014	8704	18014
		Panel b) Second reform sample		
Mother's PL	0.026**	-0.023		
	(0.010)	(0.014)		
Father's PL	0.034	0.036		
	(0.022)	(0.030)		
REFORM	[0.011]	[0.010]	-0.041	0.102
	[0.012]	[0.014]	(0.164)	(0.236)
Controls	Yes	Yes	Yes	Yes
R2	0.047	0.683	0.688	0.683
F	29.497	41.427	25.744	37.474
N	8301	16602	8558	16602

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table 9 The effect of parental leave on fathers' earnings at child age 4.

	CS	FE	DD	DDD
		Panel a) First reform sample		
Mother's PL	0.013	0.000		
	(0.007)	(0.011)		
Father's PL	0.035	-0.076**		
	(0.019)	(0.027)		
REFORM	[0.000]	[-0.025]	-0.256	-0.186
	[0.011]	[0.018]	(0.165)	(0.221)
Controls	Yes	Yes	Yes	Yes
R2	0.058	0.706	0.706	0.706
F	39.912	11.074	10.795	11.139
N	9007	18014	8704	18014
		Panel b) Second reform sample		
Mother's PL	0.007	0.005		
	(0.008)	(0.012)		
Father's PL	0.010	-0.075**		
	(0.020)	(0.026)		
REFORM	[0.003]	[-0.022]	-0.050	-0.074
	[0.007]	[0.014]	(0.138)	(0.206)
Controls	Yes	Yes	Yes	Yes
R2	0.047	0.714	0.731	0.713
F	25.454	3.860	2.125	3.031
N	8301	16602	8558	16602

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

5.2 Robustness: other specifications

In the main analysis above, the dependent variable is defined as log(earnings+1) to include also individuals who do not participate in the labor market. As discussed above, this is not unproblematic and tables A3 and A4 in Appendix show alternative specifications for the effect of parental leave/the reforms on the probability of having nonzero earnings (the extensive margin) and on log earnings among those with earnings>0, using the FE or DDD models.

The effect of parental leave on the participation decision is mostly not statistically significant, but the effect on log earnings among those with earnings >0 follow the same pattern as above – a negative effect of own parental leave and, for mothers, a positive effect of spousal leave in the second reform sample. The magnitudes of the effects are, as expected, smaller since now zero observations are excluded and part of the effect in the main analysis above was driven by individuals with zero earnings. Again, the DDD model returns only imprecisely estimated effects. [13]

6 Extensions

6.1 Heterogeneous effects

Usually, career interruptions are believed to be more harmful for individuals in occupations requiring a high level of human capital input. Therefore, we may hypothesize that both own and spousal parental leave is more important for parents with a high level of education. Also, as we saw above, the responsiveness to the reforms differed slightly between groups. However, estimating the models (FE/DDD) separately for subgroups with different maternal and paternal levels of education yields mostly imprecisely estimated effects that are not significantly different between the groups. This is most likely because of the smaller sample sizes in the FE case.

6.2 The effect of non-holiday parental leave

If there is an effect of fathers' leave on mothers' labor market behavior, one might hypothesize that this effect should differ depending on the timing of this leave. In particular, the great flexibility of the Swedish parental leave (remember that the days can be used until the child turns eight years old) also means that parents can use parental leave instead of ordinary vacation, for example during summertime or around Christmas. Such parental leave is potentially less helpful for mothers' careers than parental leave used when the other spouse is working.

Table 10 shows the effect of non-holiday parental leave, which is defined as parental leave excluding leave in June, July or August. This is estimated using the fixed-effects specification (model 2). Indeed, and in line with the hypothesis, non-holiday parental leave seems to have a larger negative effect on own earnings than summertime leave, and father's non-holiday leave has a larger positive effect on maternal earnings than leave including summertime leave. For example, fathers' non-holiday leave increases maternal earnings by almost 10 percent in the first reform sample (compared to 6.7 percent for all types of parental leave; see Table 8).

[13] In addition, using the models above (eq. 1-4) with earnings in levels (SEK, including zeroes) instead of in logs yields similar results as when earnings in logs are used, which indicates that the results are not sensitive to the logarithmic transformation.

Table 10 The effect of non-holiday parental leave

	FE: Effects on lnE mothers	FE: Effects on lnE fathers
	Panel a) First reform sample	
Mother's PL	-0.056***	0.002
	(0.017)	(0.015)
Father's PL	0.098**	-0.092**
	(0.037)	(0.035)
Controls	Yes	Yes
REFORM	[0.081]	[-0.033]
	[0.030]	[0.022]
R2	0.656	0.706
F	46.074	11.091
N	18014	18014
	Panel b) Second reform sample	
Mother's PL	-0.030	0.005
	(0.018)	(0.016)
Father's PL	0.057	-0.088**
	(0.036)	(0.032)
Controls	Yes	Yes
REFORM	[0.016]	[-0.026]
	[0.018]	[0.017]
R2	0.683	0.714
F	41.581	3.785
N	16602	16602

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

6.3 Other outcomes: fertility and marital/cohabitation status

A more equally shared parental leave could affect other outcomes than earnings. For example, previous studies have found that the amount of gender equality within a family may affect (increase) both fertility and marital happiness (Cooke 2004; Coltrane, 2000; De Laat and Sevilla Sanz, 2006; Nilsson, 2008; Oláh; 2003; Sacerdote and Feyrer, 2008; Torr and Short, 2004).

Tables 11 and 12 below show the effects of parental leave/the reforms on fertility and cohabitant/marital status, at child age 4. Since we focus on first-born children, the number of siblings is always zero before the child is born; hence, in the siblings regression we cannot make within family comparisons over time. Therefore, results are shown for the cross-sectional model and for a “horizontal” DD-model, where the number of siblings is compared across cohort and month-of-birth (instead of across time and month of birth in the standard DD-model). For the regressions on cohabitant/marital status, the FE and DDD-specifications are used.

Clearly, and in line with previous studies, both mothers’ and fathers’ parental leave have positive effects on fertility and the probability of cohabiting and being married. The coefficients in the cross-sectional and fixed-effects models are always statistically significant and very close in magnitude over time (first versus second reform sample). This suggests ambiguous expected effects of the first reform since it decreased mothers' leave while increasing fathers' leave, and positive effects of the second reform. Turning to the DD/DDD models, the results are again imprecisely estimated, but the point estimates for fertility are quite close to the predicted effects as suggested by the CS model.

Table 11 Effects on fertility (no. of younger siblings)

	CS	**DD-variant**
	Panel a) First reform sample	
Mother's PL	0.057***	
	(0.001)	
Father's PL	0.065***	
	(0.002)	
REFORM	[-0.028]	-0.022
	[0.020]	(0.022)
Controls	Yes	Yes
R2	0.328	0.032
F	382.805	27.570
N	9007	9007
	Panel b) Second reform sample	
Mother's PL	0.052***	
	(0.001)	
Father's PL	0.055***	
	(0.002)	
REFORM	[0.019]	0.011
	[0.017]	(0.023)
Controls	Yes	Yes
R2	0.272	0.057
F	272.253	45.510
N	8301	8301

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table 12 Effects on cohabitant/marital status

	Prob(cohabiting)		Prob(married)	
	FE	**DDD**	**FE**	**DDD**
		Panel a) First reform sample		
Mother's PL	0.010***		0.010***	
	(0.001)		(0.001)	
Father's PL	0.016***		0.018***	
	(0.002)		(0.003)	
REFORM	[-0.003]	-0.016	[-0.003]	-0.008
	[0.004]	(0.021)	[0.004]	(0.025)
Controls	Yes	Yes	Yes	Yes
R2	0.878	0.875	0.794	0.791
F	2017.344	1653.074	179.212	160.217
N	18014	18014	18014	18014
		Panel b) Second reform sample		
Mother's PL	0.008***		0.009***	
	(0.001)		(0.001)	
Father's PL	0.016***		0.020***	
	(0.002)		(0.003)	
REFORM	[0.005]	-0.011	[0.007]	-0.028
	[0.003]	(0.019)	[0.004]	(0.026)
Controls	Yes	Yes	Yes	Yes
R2	0.905	0.903	0.809	0.806
F	2897.678	2408.354	151.733	135.382
N	16602	16602	16602	16602

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

7 Concluding remarks

This paper investigates the effect of parental leave on earnings. In contrast to most previous studies, not only own but also spousal parental leave is considered, under the hypothesis that spousal help in child care may feed back onto each individual's labor market behavior.

Using a fixed effects model to account for time-constant unobserved heterogeneity, the results show that own parental leave is associated with earnings reductions of 4.5 percent for mothers and 7.5 percent for fathers. In terms of sign, this is in line with previous studies. The size of the effects is much larger than in previous studies, partly because the focus here is on annual earnings (which also reflect hours worked) as compared to wages, which is mostly used in other studies.

For mothers, also spousal parental leave is important for future earnings. Each month that the father stays on parental leave increases maternal earnings by 6.7 percent, which is an even larger effect than the mother's own leave. This suggests that paternal (lack of) involvement in child care and parental leave could be one factor behind the remaining, unexplained earnings gap. Among fathers, there is no effect of spousal parental leave on earnings. Even larger effects of fathers' leave on maternal earnings can be found if we restrict focus to "non-holiday" parental leave, i.e. parental leave excluding leave during the summer (June, July, or August). Such parental leave may be a better measure of spousal help than parental leave during summertime (when both spouses may be at home simultaneously because of ordinary vacation).

Finally, the fixed-effects model rests on the assumption of no unobserved, time-variant heterogeneity. In particular, it assumes that parental leave is unaffected by for example income shocks. If this assumption is violated, we need some kind of exogenous variation to identify causal effects. The two daddy-month reforms in 1995 and 2002 had a strong effect on parental leave usage. Despite that, using the reforms as exogenous variation in parental leave yields only very imprecise estimates. This is most likely due to large random variation in earnings depending on child birth dates. However, the point estimates from DD and DDD models tentatively suggests effects in the same range or larger than what was found using the fixed-effects specification.

References

Albrecht, J.W, P-A Edin, M. Sundström and S. B. Vroman (1999): Career interruptions and subsequent earnings: a reexamination using Swedish data, The Journal of Human Resources, vol. 34, no. 2

Batljan, I., S. Tillander, S. Örnhall Ljung and M. Sjöström (2004): Föräldrapenning, pappornas uttag av dagar, fakta och analys, Ministry of Health and Social Affairs (Socialdepartementet)

Becker, G. S. (1991): A treatise on the family. Enlarged edition. Cambridge, MA: Harvard university press

Berggren, S. (2005): An overview of the Swedish family benefits – goals and developments [in Swedish; English summary], working papers in social insurance 2005:1, Swedish Social Insurance Agency

Browning, M. (1992): Children and household economic behavior, Journal of Economic Literature, vol. 30, no. 3

Coltrane, S. (2000): Research on household labor: modeling and measuring the social embeddedness of routine family work, Journal of Marriage and the Family, vol. 62, no. 4

Cooke, L.P. (2004): The gendered division of labor and family outcomes in Germany, Journal of Marriage and Family, vol. 66

Datta Gupta, N., N. Smith and M. Verner (2008): The impact of Nordic countries' family friendly policies on employment, wages, and children, Review of Economics of the Household, vol. 6, no. 1

Datta Gupta, N. , N. Smith and L. S. Stratton (2007): Is marriage poisonous? Are relationships taxing? An analysis of the male marital wage differential in Denmark, Southern Economic Journal, vol. 74, no. 2

Datta Gupta, N. and N. Smith (2002): Children and career interruptions: the family gap in Denmark, Economica, vol. 69
De Laat, J. and A. Sevilla Sanz (2006): Working women, men's home time and lowest low fertility, ISER working paper 2006-23, University of Essex, Colchester
Duvander, A-Z, T. Ferrarini and S. Thalberg (2005): Swedish parental leave and gender equality. Achievements and reform challenges in a European perspective, Institute for future studies, working paper 2005:11
Ekberg, J., R. Eriksson and G. Friebel (2005): Parental leave – a policy evaluation of the Swedish "daddy-month" reform, IZA discussion paper no. 1617, Institute for the Study of Labor, Bonn, Germany
Ekberg, J., R. Eriksson and G. Friebel (2004): Sharing responsibility? Short- and long-term effects of the Swedish "daddy-month" reform, working paper 3/2004, Swedish Institute for Social Research (SOFI)
Evertsson, M. and M. Nermo (2007): Changing resources and the division of housework: a longitudinal study of Swedish couples, European Sociological Review, vol. 23, no. 4
Gangl, M. and A. Ziefle (2009): Motherhood, labor force behavior, and womens careers: an empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States, Demography, vol. 46, no. 2
Gershuny, J. and J.P. Robinson (1988): Historical changes in the household division of labor, Demography, vol. 25, no. 4
Gray, J. S. (1997): The fall in men's return to marriage: declining productivity effects or changing selection?, Journal of Human Resources, vol. 32, no. 3
Gronau, R. (1988): Sex-related wage differentials and women's interrupted labor careers – the chicken or the egg, Journal of Labor Economics, vol. 6, no. 3
Görlich, D. and A. De Grip (2009): Human capital depreciation during hometime, Oxford Economic Papers, vol. 61, supplementary issue 1
Halleröd, B. (2005): Sharing of housework and money among Swedish couples: do they behave rationally?, European Sociological Review, vol. 21, no. 3
Hersch, J. and L.S. Stratton (2000): Housework and wages, The Journal of Human Resources, vol. 37, no 1
Hersch, J. and L.S. Stratton (1997): Housework, fixed effects, and wages of married workers, The Journal of Human Resources, vol. 32, no 2
Hersch, J. and L.S. Stratton (1994): Housework, wages and the division of housework time for employed spouses, American Economic Review, vol. 84, no. 2
Karlsson, B. (1994a): Cirkus kring pappamånad. Kds, m och c hotar att rösta nej om inte vårdnadsbidraget går igenom, Dagens Nyheter, 27 May
Karlsson, B. (1994b): Smitare vid röstning om pappamånad, Dagens Nyheter, 2 June
Lundberg, S. and R.A. Pollak (2007): The American family and family economics, NBER working paper no. 12908, National Bureau of Economic Research
Lundberg, S (2005): Men and islands: dealing with the family in empirical labor economics, Labour Economics, vol. 12
Mincer, J. and H. Ofek (1982): Interrupted work careers: depreciation and restoration of human capital, The Journal of Human Resources, vol. 17, no. 1
Mincer, J. and S. Polachek (1974): Family investments in human capital: earnings of women, The Journal of Political Economy, vol. 82, no. 2
Mincer, J (1974): Schooling, experience, and earnings, New York: Columbia University Press

Nepomnyaschy, L. and J. Waldfogel (2007): Paternity leave and fathers' involvement with their young children, Community, Work and Family, vol. 10, no. 4

Nilsson, K. (2008): Jämställdhet, barnafödande och separationer, in Jämställdhetens pris, eds. A. Grönlund and B. Halleröd, Borea bokförlag, Umeå

Oláh, L. SZ. (2003): Gendering fertility: second births in Sweden and Hungary, Population Research and Policy Review, no. 22, no.2

Pylkkänen, E. and N. Smith (2003): Career interruptions due to parental leave: a comparative study of Denmark and Sweden, OECD Social, Employment and Migration working papers

Ruhm, C. (1998): The economic consequences of parental leave mandates: lessons from Europe, The Quarterly Journal of Economics, vol. 113, no. 1

Sacerdote, B. and J. Feyrer (2008): Will the stork return to Europe and Japan? Understanding fertility within developed nations, NBER working paper no. 14114, National Bureau of Economic Research

Skyt Nielsen, H. (2009): Causes and consequences of father's child leave: evidence from a reform of leave schemes, IZA discussion paper no. 4267, Institute for the Study of Labor, Bonn, Germany

Spence, M. (1973): Job market signaling, Quarterly Journal of Economics, vol. 87, no. 3

Stafford, F.P. and M. Sundström (1996): Time out for childcare: signalling and earnings rebound effects for men and women, Labour, vol. 10

Statistics Sweden (2000): Barn och deras familjer 1999, Demografiska rapporter 2000:2

The Swedish Government (1994): Proposition 1993/94:147: Jämställdhetspolitiken: Delad makt - delat ansvar

TT (Tidningarnas Telegrambyrå/Multi media news provider) (2001): Kontaktdagar och pappamånad klubbad i riksdagen, 22 Mars

Swedish National Agency for School Improvement (2007): The Swedish pre-school class – one of a kind (pamphlet)

Swedish Social Insurance Agency (2002): Föräldrapenning, Vägledning 2002:1, version 2 [in Swedish].

Torr, B.M. and S.E. Short (2004): Second births and the second shift: a research note on gender equity and fertility, Population and Development Review, vol. 30, no.1

Wooldridge, J. (2002): Econometric analysis of cross section and panel data, Cambridge: The MIT Press

Appendix

A1 Additional tables

Table A1 The effect of the reform on PL usage, by mother's level of education

	Mother's PL: low educ.	**Father's PL: low educ.**	**Mother's PL: high school educ.**	**Father's PL: high school educ.**	**Mother's PL: university educ.**	**Father's PL: university educ.**
			Panel a) First reform sample			
REFORM	-40.484	-6.606	-31.210**	12.231*	-7.434	12.772
	(29.238)	(13.599)	(12.074)	(4.930)	(17.149)	(7.954)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.897	0.557	0.908	0.596	0.904	0.627
F	408.557	15.586	2386.775	126.709	1081.300	89.595
N	2114	2114	10754	10754	5146	5146
			Panel b) Second reform sample			
REFORM	-2.458	-4.087	2.506	13.550*	1.821	6.070
	(30.748)	(15.297)	(12.454)	(5.301)	(13.616)	(7.256)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.885	0.609	0.904	0.643	0.902	0.686
F	292.824	21.625	1822.531	165.912	1289.197	182.773
N	1730	1730	8606	8606	6266	6266

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table A2 The effect of the reform on PL usage, by father's level of education

	Mother's PL: low educ.	**Father's PL: low educ.**	**Mother's PL: high school educ.**	**Father's PL: high school educ.**	**Mother's PL: university educ.**	**Father's PL: university educ.**
			Panel a) First reform sample			
REFORM	-2.510	13.989	-35.413**	9.492	-16.395	9.721
	(27.176)	(12.088)	(11.905)	(5.106)	(18.059)	(7.803)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.896	0.574	0.909	0.594	0.903	0.632
F	476.991	22.888	2390.505	121.930	1016.760	86.887
N	2470	2470	10654	10654	4890	4890
			Panel b) Second reform sample			
REFORM	4.451	7.799	2.431	6.844	1.156	12.650
	(27.334)	(12.110)	(11.468)	(5.389)	(16.488)	(8.178)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.900	0.616	0.905	0.642	0.894	0.695
F	371.289	27.823	2122.051	191.315	911.531	151.987
N	1826	1826	9846	9846	4930	4930

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table A3 Robustness, mothers

	Prob(earnings>0)		LnE given earnings>0	
	FE	**DDD**	**FE**	**DDD**
		Panel a) First reform sample		
Mother's PL	-0.003*		-0.017***	
	(0.001)		(0.005)	
Father's PL	0.004		0.024	
	(0.003)		(0.014)	
REFORM	[0.004]	0.016	[0.022]	-0.031
	[0.002]	(0.022)	[0.009]	(0.096)
Controls	Yes	Yes	Yes	Yes
R2	0.609	0.609	0.677	0.677
F	21.544	19.976	45.538	41.773
N	18014	18014	16306	16306
		Panel b) Second reform sample		
Mother's PL	-0.001		-0.011*	
	(0.001)		(0.005)	
Father's PL	-0.002		0.061***	
	(0.003)		(0.011)	
REFORM	[-0.001]	0.007	[0.018]	0.032
	[0.001]	(0.021)	[0.010]	(0.092)
Controls	Yes	Yes	Yes	Yes
R2	0.639	0.640	0.689	0.684
F	18.358	16.724	46.653	39.258
N	16602	16602	15239	15239

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

Table A4 Robustness, fathers

	Prob(earnings>0)		LnE given earnings>0	
	FE	**DDD**	**FE**	**DDD**
		Panel a) First reform sample		
Mother's PL	0.000		-0.001	
	(0.001)		(0.003)	
Father's PL	-0.004		-0.028**	
	(0.002)		(0.009)	
REFORM	[-0.002]	-0.023	[-0.009]	0.074
	[0.001]	(0.019)	[0.006]	(0.065)
Controls	Yes	Yes	Yes	Yes
R2	0.658	0.659	0.769	0.769
F	1.691	2.227	72.935	68.220
N	18014	18014	16530	16530
		Panel b) Second reform sample		
Mother's PL	0.001		-0.004	
	(0.001)		(0.003)	
Father's PL	-0.003		-0.045***	
	(0.002)		(0.008)	
REFORM	[-0.001]	-0.005	[-0.014]	-0.021
	[0.001]	(0.017)	[0.007]	(0.061)
Controls	Yes	Yes	Yes	Yes
R2	0.676	0.675	0.748	0.746
F	1.199	0.770	41.415	35.490
N	16602	16602	15570	15570

Notes: Significance levels: * 10 %, ** 5%, *** 1%. Standard errors in parentheses, clustered on family.

A2 The timing of variable collection

Figure A1 shows the timing of variable collection. All variables are collected at two points in time: one year before the birth of the child (for notational convenience this is called t=0 although it in practice means t=-1) and also at child age four (t=4). However, as is clear from the picture, this is *average* child ages. Since the variables are measured the 31st of December each year, this will mean that children born in January will on average be one month younger than children born in December when the variables are collected.

The parental leave variables are measured as the cumulative amount of parental leave up to child age three. The motivation is that it is not very interesting to estimate the direct effect of parental leave today on earnings today. Rather, the interesting relationship is that between early parental leave on future earnings.

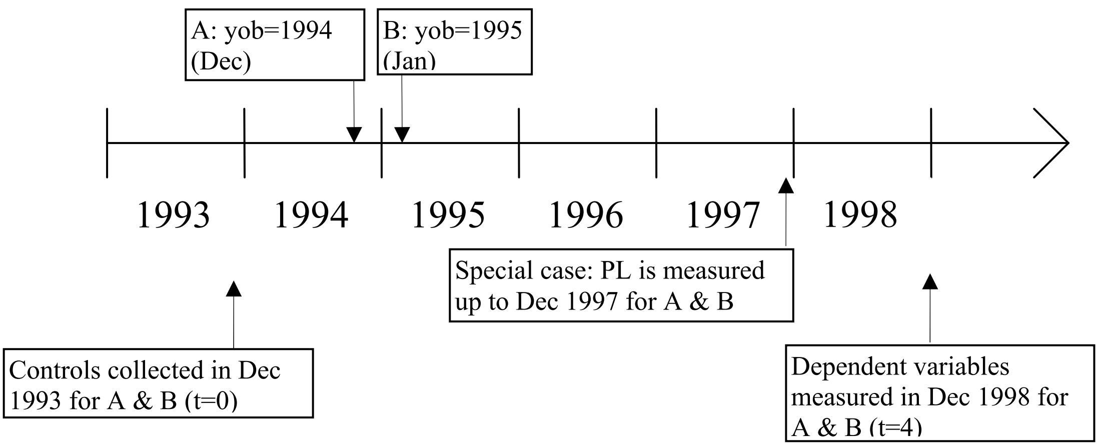

Figure A1 The timing of variable collection: example for reform cohort, first reform sample

A3 Details of the parental leave benefits over the years

Period	SGI days	% of income reimbursed	"Roof" of yearly income (SEK)	Max SEK/day, SGI days	Max SEK day if SGI=0	Flat rate days	SEK/dat, flat rate days
1990	360	90	222750	549	60	90	60
1991	360	90	241500	595	60	90	60
1992	360	90	252750	623	60	90	60
1993	360	90	258000	636	60	90	60
1994[a]	360	90	264000	651	64	90/0	60/0
1995[b]	360	80	267750	587	60	90	60
1996[c]	360	75	271500	558	60	90	60
1997	360	75	272250	559	60	90	60
1998	360	80	273000	598	60	90	60
1999	360	80	273000	598	60	90	60
2000	360	80	274500	602	60	90	60
2001	360	80	276750	607	60	90	60
2002[d]	390	80	284250	623	120	90	60
2003	390	80	289500	635	150	90	60
2004	390	80	294750	646	180	90	60
2005	390	80	295500	648	180	90	60
2006 (to June 30)	390	80	297750	653	180	90	60
2006 (from July 1)	390	80	397000	870	180	90	180
2007	390	80	398567	874	180	90	180
2008	390	80	397700	872	180	90	180
2009	390	80	415160	910	180	90	180

Notes: a) During the second half of 1994, the flat rate days were temporarily abolished for children >1 year old.

b) The first "daddy month" was introduced for children born after the 1st of january, 1995. During the 30 days set aside for each parent (the daddy month), the reimbursement level for the SGI days was still 90% of previous income.

c) During the 30 days set aside for each parent (the daddy month), the reimbursement level for the SGI days was still 85% of previous income.

d) The second "daddy month" was introduced for children born after the 1st of january, 2002.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2009:20 Böhlmark Anders, Oskar Nordström Skans and Olof Åslund ”Invandringsålderns betydelse för social och ekonomisk integration”

2009:21 Sibbmark Kristina ”Arbetsmarknadspolitisk översikt 2008”

2009:22 Eliason Marcus ”Inkomster efter en jobbförlust: betydelsen av familjen och trygghets–systemet”

2009:23 Bennmarker Helge, Erik Grönqvist and Björn Öckert ”Betalt efter resultat: utvärdering av försöksverksamhet med privata arbetsförmedlingar”

2009:24 Hensvik Lena, Oskar Nordström Skans and Olof Åslund ”Sådan chef, sådan anställd? – Rekryteringsmönster hos invandrade och infödda chefer”

2010:1 Hägglund Pathric ”Rehabiliteringskedjans effekter på sjukskrivningstiderna”

2010:2 Liljeberg Linus and Martin Lundin ”Jobbnätet ger jobb: effekter av intensifierade arbetsförmedlingsinsatser för att bryta långtidsarbetslöshet”

2010:3 Martinson Sara ”Vad var det som gick snett? En analys av lärlingsplatser för ung-domar”

2010:4 Nordström Skans Oskar och Olof Åslund ”Etnisk segregation i storstäderna – bostadsområden, arbetsplatser, skolor och familjebildning 1985–2006”

2010:5 Johansson Elly-Ann ”Effekten av delad föräldraledighet på kvinnors löner”

Working papers

2009:21 Åslund Olof, Anders Böhlmark and Oskar Nordström Skans “Age at migration and social integration”

2009:22 Arni Patrick, Rafael Lalive and Jan C. van Ours ”How effective are unemployment benefit sanctions? Looking beyond unemployment exit”

2009:23 Bennmarker Helge, Erik Grönqvist and Björn Öckert “Effects of outsourcing employ–ment services: evidence from a randomized experiment”

2009:24 Åslund Olof, Lena Hensvik and Oskar Nordström Skans “Seeking similarity: how immigrants and natives manage at the labor market”

2009:25 Karlsson Maria, Eva Cantoni and Xavier de Luna “Local polynomial regression with truncated or censored response”

2009:26 Caliendo Marco “Income support systems, labor market policies and labor supply: the German experience”

2009:27 Brewer Mike “How do income-support systems in the UK affect labour force participation?”

2009:28 Gautier Pieter A. and Bas van der Klaauw “Institutions and labor market outcomes in the Netherlands”

2009:29 Brugiavini Agar “Welfare reforms and labour supply in Italy”

2009:30 Forslund Anders “Labour supply incentives, income support systems and taxes in Sweden”

2009:31 Võrk Andres “Labour supply incentives and income support systems in Estonia”

2009:32 Forslund Anders and Peter Fredriksson “Income support systems, labour supply incentives and employment – some cross-country evidence”

2010:1 Ferraci Marc, Grégory Jolivet and Gerard J. van den Berg “Treatment evaluation in the case of interactions within markets”

2010:2 de Luna Xavier, Anders Stenberg and Olle Westerlund “Can adult education delay retirement from the labour market?”

2010:3 Olsson Martin and Peter Skogman Thoursie “Insured by the partner?”

2010:4 Johansson Elly-Ann “The effect of own and spousal parental leave on earnings”

Dissertation series

2009:1 Lindahl Erica “Empirical studies of public policies within the primary school and the sickness insurance”

2009:2 Grönqvist Hans “Essays in labor and demographic economics”

2009:3 Vikström Johan “Incentives and norms in social insurance: applications, indentifica-tions and inference”

2009:4 Nilsson Peter “Essays on social interactions and the long-term effects of early-life conditions”

2010:1 Johansson Elly-Ann “Essays on schooling, gender, and parental leave”