Parental Leave and Medium-Run Cognitive Child Outcomes: Quasi-Experimental Evidence from a Large Parental Leave Reform

Natalia Danzer
Ifo Institute Leibniz Institute for Economic Research at the University of Munich and IZA Bonn
danzer@ifo.de

Victor Lavy
Hebrew University,
University of Warwick and NBER
msvictor@huji.ac.il

Preliminary version. Please, do not cite.

February 1, 2012

Abstract

This paper investigates the question whether medium-term cognitive child outcomes are affected by the duration of maternity leave, i.e. by the time mothers spend at home with their newborn before returning to work. Employing a difference-in-difference approach, this paper evaluates an unanticipated reform in Austria which extended the maximum duration of paid and job protected parental leave from 12 to 24 months for children born on July 1, 1990 or later. The empirical analysis is based on test scores from two waves of the Austrian PISA test covering the birth cohorts 1990 and 1987. While the results based on the pooled sample suggest no significant overall impact of the extended parental leave mandate on standardized test scores at age 15, the subgroup of boys of highly educated mothers seem to have benefited.

JEL-codes: J13, J24, J22

KEYWORDS: parental leave, maternal employment, difference-in-differences, human capital, child development, cognitive skills

* Contact: danzer@ifo.de; ifo Institute Leibniz Institute for Economic Research at the University of Munich, Department of Social Policy and Labour Markets, Poschingerstr. 5, 81679 Munich, Germany

We are deeply grateful to Jonathan Wadsworth for valuable comments and discussions. We also benefited from helpful feedback from René Böheim, Andrew Clark, Alexander M. Danzer, Peter Dolton, Analía Schlosser, Uta Schönberg, and Anna Vignoles.

1 Introduction and Research Question

This paper investigates the question whether prolonged paid and job protected parental leave has effects on the cognitive development through the increase in the maternal time that can be devoted to child-rearing.

the regulations in the FMLA apply to only about half of the female workforce; parental leave allowances for the other half are determined in individual employer regulations.[3] In contrast, most European countries grant much longer durations of job protected leave (some of which is even compulsory) and mothers or fathers on leave receive partial or full income compensation (for an overview, see Neyer 2003). Since the length of the granted parental leave is relevant for the return-to-work decision, these cross-country differences in legislation help to explain why new mothers in some countries return to work much sooner and spend less time at home with their child compared to mothers in other countries (e.g. Ruhm 2000 for a study on 16 European countries; Tanaka 2005 for an analysis of 18 OECD countries).

Overall, previous empirical studies in psychology and economics have produced mixed evidence regarding the impact of early maternal employment on child outcomes. If anything, the majority of studies seem to support the hypothesis that the labour force *first* year of life has potentially adverse effects on their subsequent development (see Ruhm 2008 and studies cited therein). In addition, there is some indication for heterogeneous effects across subgroups: children with a higher socio-economic background are potentially more likely to be *negatively* affected by maternal employment, while children from low income families might *benefit* if maternal employment improves the income situation of the household (Currie 2003). Hence, depending on the specific design of the laws regarding the length of granted leave, the income replacement level during the leave period as well as the medium- to long-term labour market consequences for the mother, parental leave mandates might affect child outcomes through time effects (more maternal time investments) and potentially through income effects (if household income is reduced due to foregone wage earnings of the mother in the short-run and potentially in the long-run).

The fundamental challenge of these empirical assessments is the non-trivial identification of the causal effect of early maternal employment on child development. Maternal employment, fertility behaviour and the timing the labour market re-entry after childbirth are choice variables and might be driven by unobserved mother or child characteristics (e.g. ability, fertility and work preferences, role models, regional

[3] A more detailed description and analysis of the US Parental Leave regulations and differences across single states is given by Han, Ruhm, Waldfogel (2009).

differences in availability and costs of alternative child care). If particular types of women (e.g., higher ability women) return to work sooner than others, differences in child outcomes between these groups of mothers might reflect differences in maternal characteristics and intergenerational transmission of ability rather than the causal effect of maternal employment. Reverse causality might be an additional challenge, since certain health conditions of a child are likely to impede its

long as it is not clear which factors lead some women to return to work sooner than others, empirical analyses will only provide statistical associations and one must be careful not to draw conclusions about causality.

Until recently, the majority of studies on this topic tried to tackle the endogeneity problem of the maternal return-to-work decision (omitted variable bias) by including as many potentially relevant control variables as possible (e.g., pre-birth characteristics, the omission of which could confound the analysis), by estimating family fixed effects models and comparing sibling differences, by implementing propensity score matching or by employing instrumental variable techniques (among the first to follow this approach were Blau and Grossberg (1992); for an overview of previous studies and methods, see Currie 2003; Almond and Currie 2010; Hill, Waldfogel, Brooks-Gunn and Han 2005). However, as Currie (2003) notes, each of these empirical approaches has severe limitations and any inference and conclusions drawn from single studies have to be put in specific context and compared to results using other methods. None of these methods (except for the instrumental variable techniques if strong and compelling instruments are available), can convincingly solve the self-selection into maternal employment problem or the potential reverse causality from child needs.

A very recent strand of the literature has tried to address the identification problem by employing quasi-experimental methods. These exploit exogenous changes in maternal employment caused by reforms in parental leave provisions in several countries with different institutional settings (Baker and Milligan (2010a, 2010b) focus on Canada, Carneiro, Løken and Salvanes (2010) on Norway, Dustmann and Schönberg (2010) on Germany, Liu and Nordstrom Skans (2010) on Sweden, Würtz Rasmussen (2010) on Denmark).

This study complements these quasi-experimental studies by analysing the effects of a substantial exogenous change in the duration of maternal time at home on medium-term cognitive child outcomes in Austria. The exogenous variation was induced by an unexpected and unanticipated policy reform that extended the maximum duration of job protected and paid parental leave by twelve months for all eligible mothers giving birth on July 1, 1990 or afterwards. Employed women having a child before this cut-off date were only eligible for job protected and paid parental leave until

The reform had a strong impact on the time new mothers stayed at home with their children before returning to work since (a) female labour force participation in Austria in 1990 was already comparatively high[4], (b) most employed women generally satisfied the eligibility criteria, (c) take-up rates were extremely high and (d) most mothers exhausted the full duration of their leave entitlements (about 80 percent of mothers) (Lalive and Zweimüller 2009; Lalive, Schlosser, Steinhauer and Zweimüller 2010). However, although the reform caused mothers to substantially delay their return to work in the short-run, it did not adversely affect medium- or long-run employment and earnings of mothers (Lalive and Zweimüller 2009; Lalive, Schlosser, Steinhauer and Zweimüller 2010).

The aim of this paper is to assess the reduced form or intention-to-treat effect of this twelve-months expansion of paid and job-protected parental leave for mothers on the cognitive skills of their children at age 15, measured by test scores from standardized assessments in mathematics, reading and scientific literacy from the international PISA study (Programme for International Student Assessment). The main empirical strategy is based on a Difference-in-Differences (DID) design which exploits the variation in the duration of parental leave created by the specific cut-off date of the reform. Specifically, differences in average test scores of children born shortly before and shortly after the reform (born in May/June 1990 versus July/August 1990 respectively) are compared with the test score differences in a control year in which there was no reform (children born in May/June 1987 versus July/August 1987). The inclusion of an additional pre-reform control year is motivated by the fact that outcome comparisons across birth months within a given year could be confounded by season of birth or simple age effects (older children being more potentially advantaged at any given test date; see, for instance, Bedard and Dhuey 2006). However, as long as these

[4] More than 70 percent among 20 to 40 year old women on average and even more than 80 percent for women with post-secondary education (see also Section 4.2).

underlying seasonality or age effects are constant over time, the DID approach will difference these out and estimate the true effect of the reform.

This study contributes to the existing literature in several ways: First, it adds to the few pieces of evidence on the causal relationship between maternal employment in ognitive development identified through exogenous reductions in maternal employment. Second, the analysis sheds light on this cognitive effect in a country where most non-parental child care of under three-year-olds is provided informally by grandparents or other relatives instead of formal day care centres (in contrast to countries like Sweden and Denmark where children participate in formal care-centres already at very young ages). Against this background, the reform most likely caused a replacement of informal care through maternal care, which might have different implications than switching from formal to maternal care (the possible role of different types of child care is discussed in more detail in Section 2.2). Third, the length of the paid extension, i.e. 12 months, is much larger than in the above mentioned comparable studies using quasi-experimental designs and might thus have a stronger impact on child outcomes. Moreover, only a few papers have assessed the effect of maternal care during the *second* year of life. Fourth, in contrast to the study for Germany by Dustmann and Schönberg (2010) which comes closest to the Austrian case in terms of cultural and institutional background, this analysis contains information on parental background and can thus distinguish between heterogeneous effects across subgroups. Fifth, to the best of my knowledge, this is the first empirical evaluation of the causal link between parental leave and child outcomes in Austria.

The empirical analysis produces several results: The overall effect of prolonged parental leave on test scores for the pooled sample of all children is close to zero and statistically insignificant. The subgroup analyses by maternal education and child gender yield inconclusive results. Although the main DID estimations indicate a strong and positive effect for children (especially for boys) of highly educated mothers a critical sensitivity analysis casts doubts on the robustness of these findings. Unfortunately, data limitations prevent a final assessment of whether these ambiguous results are caused by a violation of the common trend assumption. In a future extension of the paper, we plan to repeat the empirical analysis using administrative data on school outcomes.

The paper is organized as follows: the next section introduces the underlying theoretical framework and reviews findings from previous empirical studies aiming at identifying the causal effect of early maternal employment on child outcomes in quasi-experimental approaches. This is followed by an overview of the institutional background in Section 3. The section includes details of the Austrian reform as well as a summary of findings from previous studies which evaluated this reform with respect to labour market and fertility outcomes. Section 4 explains the identification strategy and discusses critical assumptions and empirical challenges. Section 5 introduces the data set and describes the included outcome and control variables. The results of the main specification are presented and discussed in Section 6, while Section 7 contains several robustness checks and sensitivity analyses. Section 8 concludes with a critical and comparative summary of the empirical findings.

2 Parental Leave, Maternal Employment and Child Development

2.1 The role of maternal employment in the cognitive ability production function

Since the formation of human capital during childhood is a very complex process it is helpful to structure the discussion and analysis of potential effects of maternal employment on child outcomes along the lines of an underlying cognitive ability production function of the following type, where Y_{it} denotes a measure of cognitive ability of child *i* at age *t* (a similar formulation has been used, for instance, by Bernal 2008, Bernal and Keane 2011, and Dustmann and Schönberg 2010).

$$\boldsymbol{Y_{it} = Y_{it}(T_{it}, G_{it}, C_{it}, F_{it}, P_{it}, \omega_i)} \tag{4.1}$$

According to this framework the cognitive ability of child *i* at age *t* is determined by several input factors, namely (T) maternal (parental) time investment up through age *t*, (G) inputs in the form of market-purchased goods and services other than non-parental child care (depending on income of parents; examples are quality of housing, additional educational material, nutrition, health expenditure), (C) time investment through non-parental caregivers (i.e. time in non-parental child care), (F) any direct effect of family composition, e.g. number of siblings (interaction between

siblings; quantity-quality trade-off), birth order[5], time intervals between siblings[6], (P) public investments in children and child development (e.g., early child development programmes, public child care facilities and schools, state child health programmes and health insurance) and ($\boldsymbol{\omega}$) an idiosyncratic ability endowment, e.g., through intergenerational transmission of genes. As the function differentiates between different child ages it allows for varying effects of certain inputs at particular stages of child development.

Obviously, families cannot increase all input factors at the same time due to monetary and time budget constraints. In particular, when analysing the role of maternal employment within this framework, there is a clear trade-off between maternal time investment and maternal earned income which could be used to buy market-based inputs. Furthermore, it is possible that the reduction in maternal time inputs of working mothers can be at least partly compensated by the input of other goods (e.g., health investments, better nutrition) or by higher quality time investments of other caregivers (high income parents might be able to afford better quality child care). However, it should be noted that the time-income trade-off can be mitigated to the extent that mothers receive compensating maternal leave payments while being on parental leave.

The general provision or existence as well as the amount of such payments vary substantially across countries. In Austria at the beginning of the 1990s, mothers on parental leave received a monthly flat payment of approximately one third of the median female earnings (see Section 3.1). When assessing the importance of potential income losses related to leave periods after childbirth, it is crucial to consider not only short-term income losses due to absence from work during parental leave, but also potential long-term reductions in earnings which could result from human capital depreciation and/or increased difficulties in finding an employment after a prolonged leave period that exceeded the period of legal job protection.

[5] Using a rich data set from Norway, Black, Devereux and Salvanes (2005) find that an increasing number of siblings leads to a reduction of average quality he within-siblings and that children of higher birth order are more adversely affected.

[6] A recent study from Sweden uses exogenously determined changes in spacing between births and finds that shorter spacing (within 24 months) has a detrimental effect on child outcomes (Pettersson-Lidbom and Skogman Thoursie 2009). Buckles and Munnich (2011) use instrumental variables to correct for the endogeneity of the time interval between births and find that test scores of older siblings are lower when spacing between siblings is smaller (preliminary findings based on US data from the NLSY79 Children and Young Adults Survey). In contrast to these results, an earlier study from the Netherlands finds no effect of the length of the time interval between births on cognitive child outcomes (Belmont, Stein and Zybert 1978).

Further refinements of the categories of input factors in the cognitive ability production function allow a more differentiated discussion help to explain potential heterogeneity across subgroups. For instance, Blau and Grossberg (1992) disaggregate parental time into quantity of maternal time and quality of parental time where higher quality of time also leads to better child outcomes. Interestingly, in their analysis quality of parental time

A direct effect of higher education on quality of time and child care could be explained through better access to knowledge and application of methods to foster child development, for instance. In their analysis, Blau and Grossberg (1992) take quality of parental time as given; however, it is possible that working parents endogenously adjust the quality of time with their children in response to their working hours to compensate for foregone time with children. On the other hand, if parents are stressed and exhausted after work, they may have only limited capacities to spend as much quality time with and pay as much attention to their children as they would like to or as might be beneficial for the children (Baum II 2003). This aspect of endogenous quality of parental time cannot be properly addressed in most empirical analyses due to the lack of measures of actual quality of parental time. Analogously, the unobservable *quality of non-parental child care*, i.e. the relative quality of time investment from alternative caregivers in comparison with maternal care, is also likely to affect child development (Bernal 2008; Bernal and Keane 2011). There is likely to be substantial variation in the quality of child care not only through differences across formal childcare centres, but also through quality differences across alternative forms of child care: formal (accredited) versus informal child care, informal child care by relatives versus non-relatives. It seems reasonable to believe that the availability and quality of certain types of child care is likely to play an important role as an intermediating factor for the effect of maternal employment on child outcomes as these determine the relative quality of maternal and non-maternal child care. If quality of non-parental child care is positively correlated with child care costs, then access to high quality day care centres is likely to be unequally distributed across families with different levels of income.

Certainly, a reduction in early maternal employment after childbirth through extended parental leave entitlements allows mothers to spend more time with their children than otherwise. Several potential mechanisms through which increased maternal time might positively affect child development have been put forward in the existing literature. One major channel put forward in most studies is prolonged

of life, which is claimed to lead to better health outcomes of children (see critical discussion in Baker and Milligan 2008).[7] Increased maternal care time might furthermore positively influence health outcomes of children and thus their cognitive development through better monitoring ability of their health status and more timely doctor visits (see Berger, Hill and Waldfogel 2005 and studies cited therein), through more time for healthier meal preparation and house cleaning or lower risk of injuries and infectious disease (Morrill 2011). Early maternal employment, especially when exceeding 10 hours per week, might also negatively influence the attachment of mother and child and might lead to behavioural problems of the child (Brooks-Gunn, Han and Waldfogel 2002). Moreover, it is possible that exhausting market-based work leaves mothers not only with less time, but also with less energy for stimulating and nurturing child care to foster cognitive child development (lower quality home environment for cognitive development) (Ruhm 2004; Waldfogel, Han and Brooks-Gunn 2002). Conversely, a prolonged absence from work might raise the risk of social detachment and of depressions by mothers who stay at home, which in turn lowers the quality of maternal time and might have adverse effects for the children (Baum II 2003).

It should be noted that the above arguments and derived hypotheses from this simplified framework

input constant. This assumption is less restrictive in countries where the bulk of parental child care is traditionally supplied by the mother and family life follows the male bread-winner model . However, a more sophisticated analysis could furthermore take into account the joint labour force decision of the couple as well as the role of potentially endogenous paternal time investments.

2.2 Empirical evidence on the causal effect of maternal employment and cognitive child outcomes based on changes in parental leave legislations

While the relationship between maternal employment and cognitive child development has received a lot of attention in the literature, in recent years the research

[7] However, although Baker and Milligan (2008) find a significant impact of prolonged parental leave entitlements on breastfeeding duration in Canada, most of their results do not reveal any positive health effects for children.

question has been revisited with the objective to identify the *causal effect* of maternal employment by exploiting exogenous variations in maternal employment generated by changes in statutory parental leave entitlements. Table 1 provides a summary of these papers focusing on the main features of the respective reforms and those empirical results which are most relevant for the present study. All these studies have as a common starting point an unexpected extension of the granted parental leave duration which significantly increased the time mothers stayed at home before returning to work after childbirth. Furthermore, almost all studies are able to implement either a Difference-in-Difference or a Regression Discontinuity estimation strategy as most of the reforms were implemented with a sharp cut-off date that strictly determined y to extended leave periods (only in Sweden the 90-days extension was gradually phased in over a period of three months). Against the background of the theoretical considerations regarding the potentially adverse effects of early maternal employment on child development (as well as on maternal health), it seems surprising that most of these studies find either no or only negligible effects of prolonged parental leave on cognitive child outcomes. The only exception is the analysis by Carneiro, Løken and Salvanes (2010) for Norway: based on a sample of mothers who were eligible for parental leave, the authors find significant positive effects of prolonged maternal medium-run schooling achievement, IQ measures and height. The effects seem to be even stronger for children of mothers with low educational attainment. Interestingly, the Norwegian study is unique in that it can distinguish between children of eligible and ineligible mothers. The authors demonstrate that results become *insignificant* as in the other studies when children of all mothers irrespective of their eligibility and affectedness are included in the estimation sample.[8]

The findings by Liu and Nordstrom Skans (2010) for Sweden stand in contrast to the empirical results for Norway: although the average effect for the pooled sample is statistically not significantly different from zero, it appears that children of better-educated mothers having at least some tertiary education have benefited from the reform and perform better in nationally administered tests. Hence, while in Norway the effect

[8] When running estimations on the whole population of eligible and ineligible mothers, it will be the more difficult to detect the actual effect of the reform on child outcomes the larger the group of non-eligible mothers who do not change their return-to-work behaviour in response to the policy change. This is related to the fact that the effect will be averaged over the entire population and not over the actually affected subgroup.

of maternal leave on child outcomes is positive for children from lower-educated mothers, results for Sweden imply that especially children from better-educated mothers have benefited from the reform.

In order to reconcile the seemingly contradictory evidence of zero effects and positive effects for separate subgroups it is important to pay attention to the essential differences between these studies which complicate straightforward comparisons. The analyses vary predominantly with respect to:

(1) *The affected age group of children*: Does the particular extension of parental leave allow women to stay at home longer during the first year of life (e.g., an extension from 8 to 12 months) or does it affect the period when the child is more than two years old? This point is highly relevant if the importance of maternal care varies over the different development stages of the child.

(2) *The length of the extension*: The analysed parental leave extensions vary between six weeks and 18 months. The granted length of the extension is likely to influence the additional time that mothers stay home. If there is a positive effect of maternal time on child outcomes and this effect is increasing with time input, then one would expect differential effects depending on leave duration.

(3) *The measure of cognitive development and age at its measurement*: While some studies focus on short-run effects measured before the first birthday (parent-reported assessments or psychological tests), others compare medium- or long-term outcomes up to age 29 (e.g., using completed educational attainment). On the one hand, it is possible that initial differences in early cognitive development might be mitigated, for instance, through special education or development programmes over the life course. On the other hand, it could be that particular inequalities in the very early stages of life persist over time or are even aggravated (cumulative effect of an early negative shock). Furthermore, different types of measures of cognitive development might generate different results if they a) capture different aspects or types of cognitive skills (verbal versus mathematical skills), b) are designed for a particular age (e.g., Number Knowledge Test for four year-olds versus high school

grade point average), or c) differ in their precision or level of aggregation (e.g., educational attainment categories versus standardized test scores).

(4) *Different institutional environments*: There are strong differences across countries (and over time) in terms of prevailing non-parental child care arrangements (formal centre-based or informal care by relatives) which determines the type of care likely to be substituted by prolonged maternal care. In fact, there might be in addition an interaction effect of the institutional environment and of the initial parental leave length (i.e., if the pre-reform leave lasted only 3 months mothers might have used different types of child care than if the initial leave lasted already 18 months, which could be related to minimum age requirements in formal day-care facilities).

(5) *The type of the reforms and their indirect effects on other supposedly relevant determinants of child outcomes like income and fertility*: e.g., does the reform expand the duration of fully, partly and/or unpaid leave? These indirect effects could alter the (opportunity) costs of children and also enhance or change the quantity-quality trade-off.

(6) *The precision of the data and the estimations*: Can eligible mothers be identified (only possible in the Norwegian study) or can children be linked to parents? Are the studies based on representative surveys or huge administrative datasets?

(7) *The estimation strategies*: the exact implementation of the DID and RD estimations differs across studies as do the control groups.

Table 1: Studies using changes in parental leave legislations to identify causal effects of maternal employment on child outcomes

Publication (Authors and Year)	**Country and Data Source**	**Year & Substance of PL reform**	**Assessed child outcome(s) (Short/Medium/Long run effects)**	**Main empirical method**	**Results on effect of reform on child outcomes**	**Heterogenous effects**	**Institutional background: provision of child care**
Baker and Milligan (2010a)	*Canada:* National Longitudinal Study of Children and Youth (NLSCY); about 2,000 children per cohort	December 31, 2000 Maximum duration of maternity leave benefits raised from 25 to 50 weeks (out of which 10 and respectively 35 weeks can be claimed by either mother or father) Pre-reform job protected maternity leave varied between 18 and 70 weeks across regions. Post-reform maternity leave duration increased to at least 52 weeks in all regions.	**SR effects:** Children between 7 and 24 months Parent-reported measures of temperament, motor and social development	Test of differences between average outcomes of birth cohorts born before and after the reform; (regressions based on six yearly values)	Overall small and mostly insignificant effects on the development variables	Not tested	Centre-based care for children under 12 (24) months very low (4 % and 6 %) **Mainly informal care** (about 39% and 41%)
Baker and Milligan (2010b)	*Canada:* National Longitudinal Study of Children and Youth (NLSCY); about 2,000 children per cohort	See Baker and Milligan (2010a)	**SR effects:** At ages 4 or 5 Cognitive development (e.g., Peabody Picture Vocabulary or Number Knowledge Test) Parent-reported behavioural development (e.g., hyperactivity)	See Baker and Milligan (2010a)	No significant positive effects	None (No differences by child gender or parental education)	See Baker and Milligan (2010a)

Carneiro, Løken and Salvanes (2010)	*Norway*: Administrative register data on schooling and family events and military records (linked child-parent data)	July 1, 1977 Introduction of paid PL for 18 weeks (4.5 months) with 100% income replacement as well as extension of unpaid PL from 3 to 12 months (on top of paid PL) [de facto increase in PL take-up from 8 to 12 months]	**MR & LR effects:** Dropout rates from high school (measured at age 29) College attendance (measured at age 29) IQ (males aged 18-19) Teenage pregnancy (females with birth before age 20) Height (males aged 18-19)	Non-parametric RD (1977 cohort; local linear regression) and non-parametric RD-DID (cohorts 1977 and 1975)	**Significant, positive effect** on high school graduation, college attendance and IQ (for males) for sample of eligible mothers Insignificant effects when including ineligible mothers	Yes, stronger positive effects for children from households with **lower maternal education** (less than 10 years of schooling) No differences by child gender or pre-birth household income	Extremely low enrolment rates of zero- to two-year-olds in public child care in 1977; **mainly informal** child care through relatives
Dustmann and Schönberg (2010)	*Germany*: Administrative data on public schools in three federal states (information on type of school/ track and graduation); social security data on educational attainment	Three reforms: 1979, 1986, 1992 May 1, 1979: Extension of paid+job protected PL (flat rate) from 2 to 6 months January 1, 1986: Extension of paid+job protected PL (flat rate up to month 6; means-tested from month 7 to 10) from 6 to 10 months January 1, 1992: Extension of unpaid job protected PL from 18 to 36 months (maternity leave payments up to month 18)	**MR & LR effects:** 1979 reform: wages and educational attainment at age 28 or 29 1986 reform: Graduation from academic track (before age 20) 1992 reform: Choice of school track at age 14 (8th grade) (most/medium/least academic track)	DID and TS-2SLS (RD and RD-DID as robustness check)	**No significant effects** or only extremely small positive effects Effect of expansion of 18 to 36 months even slightly negative	Not tested	Enrolment in formal day care centres low (5% for under 18-months-olds); Child care **mainly informal through grandparents** or other relatives (29 %)

Liu and Nordstrom Skans (2010)	*Sweden*: Administrative register data	August October 1988 Extension of paid PL benefits from 12 to 15 months Gradual extension by 30 days in each of three consecutive months in 1988: 1st of August/September/October	**MR effects:** Test scores from national tests during last year of compulsory school Compulsory school grades (GPA scores) scores at age 16	OLS regression of child outcomes on legal number of PL months (according to birth month of child)	Average effect on child outcomes is insignificant	**Positive effect for well-educated mothers** (some tertiary education) No differences between boys and girls	**Established public child care** system: 40-50% of children aged 1-2 in formal day care or family centres; only few children in informal care
Würtz Rasmussen (2010)	*Denmark*: Administrative register data (linked child-parent data); PISA 2000	March 26, 1984 Extension of paid PL from 14 to 20 weeks	**MR effects:** High school enrolment High school GPA Reading test scores of 15-years old children (PISA test in 2000)	RD (DID as robustness check)	No significant effects	None (No differences by child gender or parental education)	Publicly subsidized **day care system even for very young** children available

Notes: PL Parental leave; RD Regression Discontinuity; DID Difference-in-Difference; TS-2SLS Two-Sample Two-Stage-Least-Squares; IV Instrumental Variables; GPA Grade Point Average; SR/MR/LR short-run/medium-run/long-run effects

For instance, consider the seemingly contradictory findings for Norway and Sweden: one has to keep in mind that the Norwegian study focuses on a reform affecting leave taking between month 8 and 12 after childbirth and considers long-term schooling outcomes at ages 18 to 29 (e.g. high school drop-out rates). Furthermore, the predominant type of child care for children of the relevant age group was mainly informal (through relatives), suggesting a substitution effect of the reform from care by relatives to maternal care (at least relatively *more* maternal care compared to the situation before the reform in families combining both types of care). The Swedish reform affected children at a later stage of development, aged 12 to 15 months and the later observed child outcomes are test scores of national tests and grades in certain subjects during the last year of compulsory school (age 16). In contrast to the Norwegian scenario, almost half of Swedish children aged 1 to 2 during the time of the reform were in public day-care and informal child care arrangements were very uncommon. Hence, the positive effect for children of well-educated mothers in Sweden might be related to a substitution of public day care with maternal care and the possibility that well-educated mothers provide comparatively better quality care for their children than public institutions, which might not be the case for less educated mothers (Liu and Nordstrom Skans 2010).[9]

On the other hand, the authors of the study for Norway conjecture that the reform resulted in a replacement of informal care through maternal care with positive effects on all children, but seemingly more pronounced effects for children of lower educated mothers (Carneiro, Løken and Salvanes 2010). Moreover, their findings are in line with results from Bernal and Keane (2011) who show that only increased informal child care (i.e., non-centre based child care by grandparents or other relatives) has adverse effects on the cognitive development of children of single mothers in the USA, while an additional year of centre-based child care has no negative effects. Bernal and Keane (2011) relate their results to other empirical studies[10] providing evidence for negative effects of informal child care by grandparents on cognitive child development and summarize the two possible main channels for this adverse effect: (a) trained, well-

[9] Generally, parents can combine different types of child care and it is also possible that the reform helped parents to combine different types of child care more efficiently.
[10] Bernal and Keane (2011) refer to two studies from the UK from Hansen and Hawkes (2009) and Gregg, Washbrook, Propper and Burgess (2005). However, the evidence from Hansen and Hawkes (2009) regarding the effects of child care by grandparents is mixed and depends on the outcome (the study finds negative behavioural effects, especially for boys).

educated personnel in day centres could generate better cognitive stimulation for children and (b) there might be positive effects through interactions between children as well as increased educational activities fostering cognitive development. Similar results are found by Datta Gupta and Simonsen (2010) for Denmark who exploit regional differences in excess demand and waiting lists for child care: while participation in centre-based care with highly qualified personnel does not have detrimental effects on non-cognitive child outcomes in comparison with home-based care, family day care (in private homes by one child minder) has negative effects on non-cognitive outcomes, especially for boys from less educated mothers.

To sum up, the institutional background and the details of the reforms vary widely across countries and seem to play an important role for the effect of parental leave on child outcomes. In terms of institutional and cultural set-up, Austria comes

first and second birthday) is more similar to the evaluated reform in Sweden where, however, participation rates of one- to two-year-olds in formal child care are very high. Nevertheless, the reform in Sweden led only to a three months extension of parental leave, while in Austria the extension comprised 12 months. This way the analysis in this paper helps to shed more light on the influence of maternal employment beyond the

able.

Furthermore, the Austrian reform is unique in that it involves an exceptionally long extension of paid parental leave.

2.3 Heterogenous effects

Some of the quasi-experimental studies mentioned in the previous section have also investigated in more detail whether maternal employment has differential impacts across population subgroups, focusing in particular on the level of educational attainment and child gender. As Table 1 indicates, there are no universally consistent patterns across these studies with respect to maternal education. While most papers do not find any differences in the impact between children of more versus less educated mothers, the Norwegian study finds more pronounced effects for children of lower educated mothers and the Swedish results indicate positive results for better educated mothers (as has been discussed in the previous section).

In terms of child gender, none of the four papers who run separate regressions for girls and boys detect any significant differences regarding the effect of prolonged maternal leave. In contrast, several previous studies from other strands of literature suggest that maternal employment has a stronger detrimental effect on boys as compared to girls, which is partly explained by the fact that boys seem to be more vulnerable to early stressors and to react more adversely to non-maternal child care (Brooks-Gunn, Han and Waldfogel 2002).[11] These negative effects are also supported by papers assessing the effect on health outcomes (see next paragraph).[12]

2.4 Maternal employment and health outcomes of children

As mentioned before in Section 2.1, maternal employment might affect cognitive development of children also through its effect on health outcomes. There is a complementary literature on the relationship between early maternal employment and health outcomes of children. These studies face the same identification problem since a return-to-work decision might be driven by unobserved factors like poor health of their children. Three recent studies specifically accounting for the non-random selection of mothers into employment by using quasi-experimental and instrumental variables approaches are Baker and Milligan (2008) for Canada as well as Gennetian, Hill, London and Lopoo (2010) and Morrill (2011) for the USA. While the Canadian study finds strong effects of prolonged parental leave on breastfeeding duration, there are generally no positive health effects on self-reported maternal or child health. In contrast to these findings are the two studies from the US: using age-based kindergarten eligibility for the youngest child as instrumental variable for maternal employment, Morrill (2011) finds that children of working mothers have a greater risk of suffering from adverse health events (overnight hospitalization, asthma episodes, injury and poisoning) and that the negative effect is twice as large for boys than for girls (the effect for the subsample of girls is even insignificant). Gennetian et al. (2010) correct for the endogenous work decision of the mother by using the experimental design of a welfare-

---

[11] In addition, there is evidence from Denmark that informal or less qualitative day care has particularly detrimental effects on boys (Datta Gupta and Simonsen 2010).

[12] General gender differences in schooling or particular types of cognitive skills (e.g., language versus mathematical skills) might be related to underlying genetic or cultural differences (Guiso, Monte, Sapienza, and Zingales 2008; Machin and Pekkarinen 2008). Furthermore, there is evidence for son preferences leading to unequal treatment of sons and daughters with potentially long-term adverse effects for girls (although this phenomenon is generally associated with developing countries, Mammen (2011) shows that fathers in the USA spend more time with sons than with daughters).

to-work programme which increased the share of mothers working without affecting household income resources (income from the welfare programme was replaced by earned income from market based work). The results imply a negative effect of maternal employment on child health outcomes in low-income families measured at age five to nine; again, this adverse effect seems stronger for boys than for girls.

3 Institutional setting and background

This section describes in detail the content and timeline of the reform of the parental leave legislation in 1990 and gives an overview of the institutional background regarding the development of maternal employment as well as child care in Austria over time.

3.1 Parental Leave in Austria and the Reform in 1990

3.1.1 The development of parental leave in Austria

The history of parental leave in Austria dates back to 1957 when working women became entitled to an unpaid, but job protected leave of up to six months on top of the paid mandatory maternity leave of 12 weeks, making Austria the first country in Europe to introduce parental leave (Neyer 2003). Since that time childbirth related leave from work in Austria comprises two parts: a mandatory maternity leave and an optional parental leave.

3.1.2 Mandatory Maternity Leave

The length of the mandatory maternity leave was raised from 12 to 16 weeks in 1974 (Hoem, Prskawetz and Neyer 2001b). More specifically, according to the Maternity Protection Act (*Mutterschutzgesetz)* employed women are not allowed to work during the last 8 weeks before the expected birth date and 8 weeks after delivery (this period extends up to 12 weeks for multiple or premature births or Caesarean sections).[13] While being on leave women receive a maternity pay (*Wochengeld*) which equals 100 percent of the average net earnings of the preceding 13 weeks. Furthermore, during pregnancy and until 4 months after delivery, women are subject to a special employment protection and cannot be dismissed by their employer. This period of

[13] Before 1974 the length of the mandatory maternity leave was only 12 weeks.

employment protection is further extended, if the mother takes parental leave immediately after the mandatory maternity leave (job protection until 4 weeks after returning to work).[14]

3.1.3 Parental leave before the reform in 1990

After the introduction of optional unpaid parental leave in 1957 this was in fact maternal leave, since fathers were not entitled to take leave important amendments to the law followed in the years 1961, 1974 and 1990. In 1961 the maximal duration of job protected parental

mothers on leave became entitled to a cash benefit as long as their household income was below a certain threshold. The level of benefits equalled the level of unemployment benefits for single mothers and half the level of unemployment benefits for married mothers.[15] To become eligible for parental leave payments women needed to be in employment subject to compulsory social insurance contributions for at least 52 weeks during the two years preceding the first birth (20 weeks during the most recent year for higher order births).[16] Furthermore, special regulations on work requirements applied for mothers becoming pregnant while being on parental leave: if their expected birth date lay within 14 weeks (3.5 months) after the expiry of the previous parental leave period then their parental leave entitlement was renewed (*automatic renewal*; see discussion in Hoem, Prskawetz and Neyer (2001a) and Lalive and Zweimüller (2009). However, this required women to become pregnant six and a half months after their last birth and which was even biologically difficult to realize and hence uncommon.

The work eligibility requirement for young mothers was subsequently reduced to 20 weeks during the last 12 months (the age criteria for young mothers was at 20 in 1974 and was raised to 25 in 1989). 1974 saw important amendments to the law regarding cash benefits: the amount of cash benefits was uncoupled from household income by granting a flat rate benefit to all mothers on parental leave (single mothers and wives of no or low income earners received a 50 percent higher assistance). A

[14] See Law on Maternity Protection (Mutterschutzgesetz) as of March 2011, http://www.jusline.at/Mutterschutzgesetz_%28MSchG%29.html

[15] The information on the parental leave in this section draws mainly on the timeline of changes in parental leave legislation provided in Austrian Family Report 1999, Volume I., chapter 12.2.3, and Volume II., chapter 3.3.2 (BMUJF 1999a, 1999b), and in Hoem et al. (2001b, Appendix C).

[16] Periods during which women received unemployment benefits are counted towards these minimum work requirements.

subgroup of particularly disadvantaged mothers became entitled to special maternity leave payments (*Sondernotstandshilfe* at first only single mothers, in 1990 married women in households with low or no income were added to this group). In 1982 farmers and self-employed mothers became eligible for up to 16 weeks of maternity flat rate transfer payments.

3.1.4 Parental leave taking of fathers

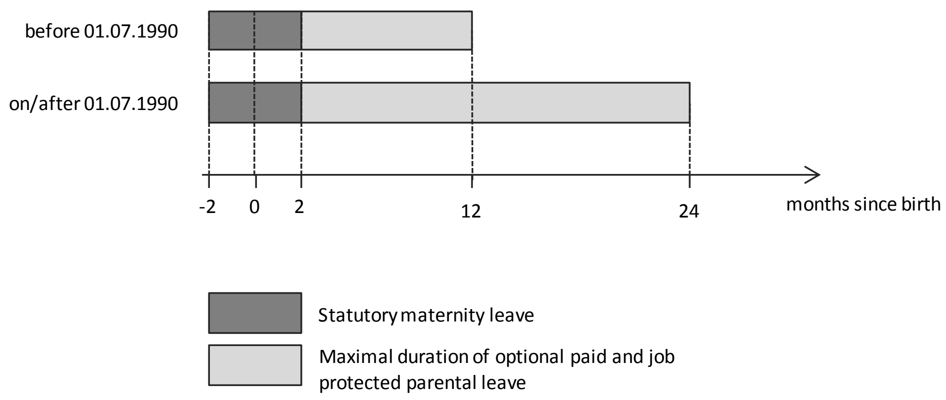
Fathers became eligible for the parental leave only as of January 1, 1990. However, their entitlement to parental leave was conditional on the mother meeting all eligibility criteria. The take-up rate of parental leave of fathers remained close to zero during the 1990s (from 0.2 percent in 1990 to 1 percent in 1997, see Table 2). These numbers demonstrate that parental leave taking was is still almost exclusively relevant to working mothers.

3.1.5 The parental leave reform in 1990

The following empirical analysis will exploit a quasi-experiment that was created by the amendment to the parental leave legislation that came in effect on July 1, 1990 (*Karenzurlaubserweiterungsgesetz, June 27, 1990, BGBl. Nr. 408/1990*). The main aspect of this reform was the extension of the maximal duration of the optional to rthday (see scheme in Figure 1). According to the law, this extension was only granted to mothers of children who were born on or after the cut-off date of July 1, 1990. There were no transition rules allowing mothers who gave birth before July 1, 1990 to benefit from the new regulations. This increase of 12 months of paid and job protected parental leave is much larger than any of the comparable reforms that took place in other countries and that have been evaluated in terms of child outcomes (see Table 1). An important side effect of this extension by 12 months was that it became easier for families to get automatic renewal of parental leave: the maximum period between previous birth and conception of a future child that would grant automatic renewal was extended by one year (from 6.5 to 18.5 months). As described in more detail further below (see Section 4.1), the reform was announced and implemented only shortly before it came into effect. This is why it was not possible for parents to adjust their fertility timing in order to take

advantage of the more generous parental leave regime (i.e. there were no anticipatory fertility effects).

Figure 1: Parental leave entitlements before and after the reform on July 1, 1990



Other changes that came into effect with this reform were the option of part-time leave between and second birthday if both parents took their leave simultaneously or up the third birthday if only one parent went on leave or the parents took the leave in turns. However, only a tiny fraction of mothers made use of this new possibility (less than 2 percent of all women on parental leave were on part-time leave in the years 1992-1994, see BMUJF 1999b, Vol. 2, Table 3.46c on page 157). Furthermore, the parental leave subsidy (*Teilzeitbeihilfe*) for farmers and self-employed amounting to half the regular flat-rate parental-leave payment was extended to the reduced parental leave subsidy became also available to employed mothers who did not meet the minimal work requirements.

In 1990, the amount of the regular flat-rate parental leave payment was about 340 Euros per month, which corresponded to 31 (40) percent of gross (net) *median* female earnings (Lalive and Zweimüller 2009).[17] In 1996 the flat-rate benefits equalled about 35 percent of the *average* monthly net female earnings (BMUJF 1999b). Since the level of parental leave benefit was independent of previous earnings, the earnings replacement rate was much lower for mothers with high pre-birth earnings than for mothers with low pre-birth earnings. Thus, the associated income loss and opportunity costs of parental leave were higher for the former than for the latter group of mothers.

---

[17] According to a study by Fuchshuber (2006), which was conducted by order of the Austrian Federal Ministry for Health and Women to investigate potential career barriers and problems for female employees, some firms (in particular larger companies) tend to offer additional benefits for families (financial benefits or fringe benefits like child care facilities) or even special parental leave or job protection arrangements beyond the legally required minimum. Unfortunately, the report does not provide information on whether these benefits were already common in 1990.

Table 2: Statistics on the development of maternity and parental leave take-up between 1985 and 1997 ETQQ2,970C q467.11 33.44 reTmBT/F23,233.56

Year	Number of women receiving mandatory maternity pay	Parental leave, 1st year	Parental leave, 2nd year	Parental leave (yearly average)	Parental leave (Dec 31)	Parental leave, Men	Parental leave, Women	Share of fathers on parental leave (%)	Number of births	Children aged 0-1	Children aged 1-2	Share of women on mand. maternity leave in comparison to all births (%)	Share of mothers on 1st year parental leave (%)
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII*)	XIII **)
1985	60,505			37,601	38,440		37,601		82,970	87,906	89,176	72.9	51.3
1986	60,412			38,132	39,031		38,132	[redacted]					

a ly 2

(a

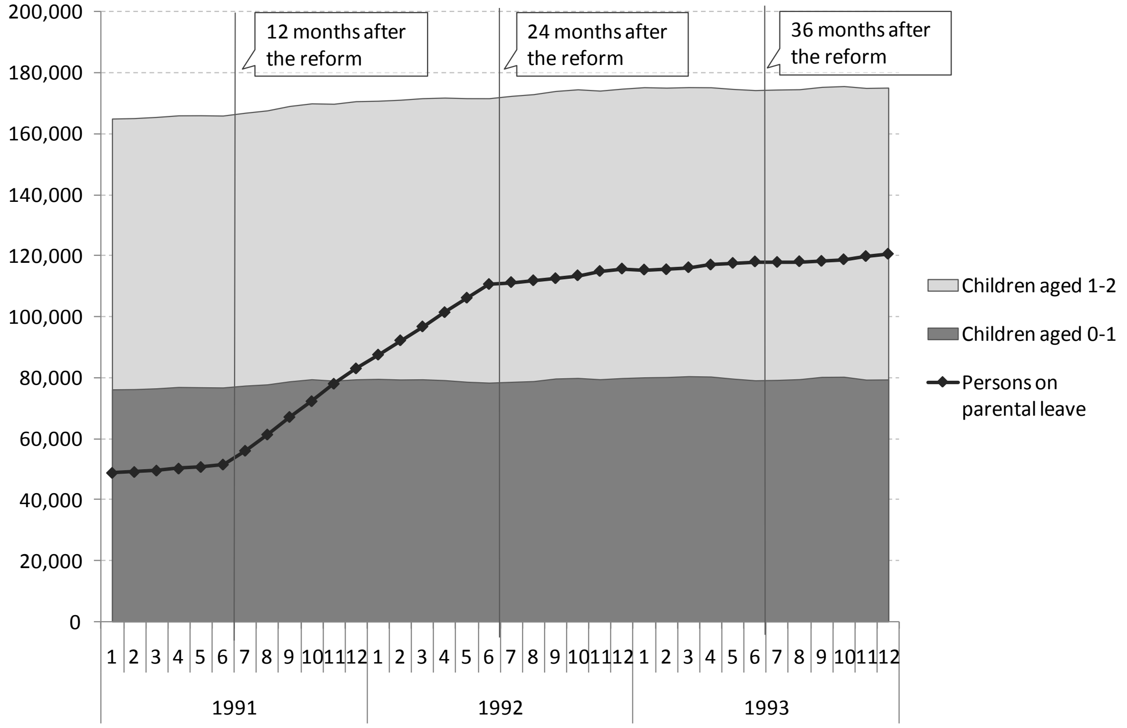
3.1.6 Share of eligible mothers among all mothers

ee

Due to the steadily increasing age at first birth (from 25 years in 1990 to 29 years in 2005; see

ot(e)4rs(-77(who(-77(we-70(e)4rel)-[(-749 e)4(li)]23(g)10iblh)-3(e44(-749fore)6( )749p)-94 a)4re

Figure 2: Numbers on take-up of parental leave as well as number of children over time showing the effect of the reform of July 1, 1990.



Source**:** Persons on Parental Leave based on statistics from AMS (AMS Public Employment Service Austria); Number of children: own calculations based on monthly birth records (Human Fertility Database). The number of children 0-1 (0-2) corresponds to the total of children born in the previous 10 (22) months.

Up to June 1991 the percentage of women on parental leave among all women who had given birth within the last year was about 65 percent (this number accounts for the fact that women appear in the statistics on parental leave for only 10 months). The first cohort of women who could remain on par

started their second leave year on July 1991 (12 months after the cut-off date). The figure demonstrates a steady inflow of mothers into the second leave year up to June 1992, which was the final, 24th month of leave for the first cohort of women who were eligible for 24 month of parental leave. Starting from July 1992 the strong inflow into the stock of persons on parental leave stopped and the share of mothers on parental leave among all mothers with up to two-year-old children stabilized again at around 65 percent (calculated using the adjusted numbers of birth in the preceding 2 years). After July 1992 the amount of women on parental leave continued to grow slightly due to the steadily rising number of eligible women (steady growth of female labour force participation and increasing age at first birth). Overall, the rise in the number of women on parental leave is approximately in line with the claim of high take-up rates of the extended leave duration: between June 1991 and June 1992 the monthly average

number of women on parental leave doubles from about 52.000 to 111.000 (the latter is the sum of mothers on first year leave *and* on second year leave ; the fact that the number more than doubles is related to the fact that there is a slight increase in number of children aged zero to two between these two points of time).

3.2 Effects on subsequent maternal labour market success and fertility

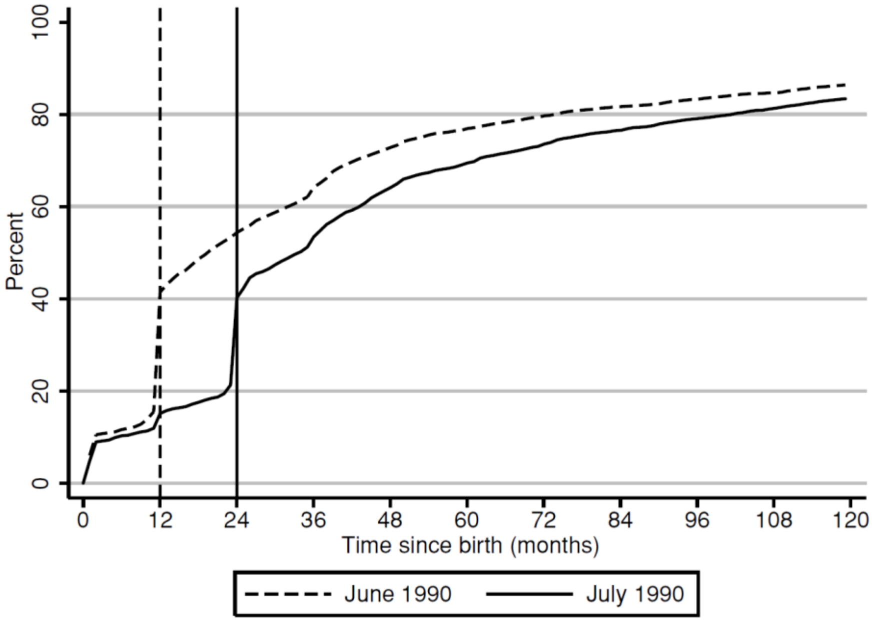
Previous studies investigating the effects of the 1990 parental leave extension have focused on outcomes and behavioural changes of new mothers, in particular their short- and medium-term labour market outcomes as well as their fertility behaviour.

3.2.1 Short-term and medium-term effects on maternal employment

The fact that the duration of the leave entitlement before and after the reform in Austria was actually binding and thereby exogenously determining the minimum length of leave take-up has been demonstrated by Lalive and Zweimüller (2009) and Lalive, Schlosser, Steinhauer and Zweimüller (2010). This result can be illustrated with Figure 3 from Lalive and Zweimüller (2009): it depicts the proportion of pre- versus post-reform mothers having returned to work at any time after giving birth to their first child. There are about ten percent of mothers who return to work immediately after the end of the mandatory maternity leave, i.e. two months after childbirth, and this pattern holds generally true before and after the reform. Both the dotted line for pre-reform mothers (June births) and the solid line for post-reform mothers (July births) reveal a jump of roughly 20 percentage points at the time when the legal leave entitlement expires (after 12 months and 24 months respectively). This reaction at the end of the parental leave shows that the duration of job-protected and paid leave is binding and determining the work decisions of a large group of mothers. Nevertheless, the majority of mothers stay out of the labour force and do not return to work immediately after the parental leave has elapsed. The effect of the reform becomes visible when considering the share of women having returned to work after 23 months: while only about 20 percent of the post-reform mothers have returned to work after 23 months, the majority (around 55 percent) of women having given birth under the old legislation have returned to work after 23 months; hence, the reduction in short-term re-entry (within 23 months) equals approximately 35 percentage points. After 24 months, i.e. after the end of the two year

parental leave, post-reform mothers are still over 10 percent less likely than pre-reform mothers to have returned to work. After three years 62 out of 100 pre-reform mothers have returned to work, whereas the same is only true for 52 out of 100 post-reform mothers. Although this gap in return-to-work rates between the two groups narrows slowly over time, post-reform mothers remain slightly less likely to have ever returned to work even after ten years (the difference amounts to three percentage points). Irrespective of the duration of the parental leave a substantial fraction of mothers do not return to work within ten years after giving birth (around 18 percent of post-reform mothers; the number for pre-reform mothers is only slightly lower).

Figure 3: Share of mothers returning to work after birth



Source: Figure 5.B. taken from Lalive and Zweimüller (2009, p. 1387)

Despite the substantial changes in return-to-work behaviour, there are no medium-term effects on alternative labour market outcomes like average number of months in employment or earnings per month (Lalive and Zweimüller 2009). Even though in the short-run post-reform mothers work significantly fewer months on average and have lower earnings than pre-reform mothers, there are no significant differences in these outcomes after ten years.

Overall, the reform has had a significant impact on the time mothers take parental leave after birth. The average duration of parental leave take-up and receiving benefits (after mandatory maternity leave) increased from 10 to 20 months (Lalive,

Schlosser, Steinhauer and Zweimüller 2010). Lalive and Zweimüller (2009, p. 1399-1340) conclude furthermore tion of their leaves (sic); that return to work is substantially delayed even after PL [parental leave] has been

generally hold for higher order births as well. As Lalive et al. (2010) demonstrate, the parental leave extension had no significant medium-term effects (after five years) on employment probability, months worked per year, daily earnings or annual incomes of mothers with first or higher order births.

Last but not least, it should be noted that mothers who gave birth after the extension of the parental leave experienced an income loss of about 3,200 Euros which, is caused almost exclusively by foregone earnings during the prolonged leave period and not by differential earnings developments after the expiration of the parental leave entitlements (Lalive et al. 2010).

3.2.2 Effects on fertility behaviour (timing and completed fertility)

Theoretical considerations as well as empirical evidence on the fertility effects of the reform are presented in the paper by Lalive and Zweimüller (2009). In their terminology, the overall impact of the reform on fertility was a combination of *current-child effect* and *future-child effect*. The current-child effect relates to changes in fertility behaviour which, in combination with the 12 months leave extension, made it biologically more feasible for couples to take advantage of the prolonged automatic renewal period in case of a new pregnancy (the automatic renewal period increased from 15.5 to 27.5 months). The authors use the term future-child effect to refer to changes in fertility due to the reduced costs associated with giving birth under the new legislation with extended job protection and benefits (the more generous parental leave regulations affect the costs-and-benefits of future births as such). Their empirical analysis of the current-child effect is based on a comparison of short- and medium-term outcomes of treated and control mothers (giving birth to their first child in July 1990 or June 1990 respectively), who do not differ in pre-birth characteristics and who face similar labour market conditions but are subject to more or less generous parental leave regulations. The empirical approach basically resembles a fuzzy Regression Discontinuity Design, since *eligibility* to the extended leave period is based on a random

assignment according to the cut-off date of childbirth, July 1, 1990 (the authors perform several sensitivity checks and provide supporting material and facts for this and other underlying identifying assumptions). The estimation of the *future child effect* (the role of extended parental leave on future births) is based on a comparison of short-term outcomes from women giving birth in June 1990 versus June 1987 (both groups are subject to the same parental leave regulations of one year for the current child, but differ with respect to the length of parental leave granted for higher order births within the next three years).

According to the analysis of Lalive and Zweimüller (2009), the reform had a strong and significant impact on fertility outcomes in the short (within three years after the previous birth) and medium-run (after ten years). The extended parental leave caused an increase in the probability of having a second child within three years after the previous birth by five percentage points (15 percent) due the *current child effect* and by seven percentage points (21 percent) due to the *future child effect*. These short run effects also translate into medium-run effects: after ten years post-reform mothers are three percent more likely to have given birth to another child than pre-reform mothers. A more detailed analysis of the timing effects reveals further interesting changes: the probability of giving birth within the first 16 months actually decreased. The overall positive effect in the three-year period is hence driven by the strong increase of the birth probability between month 17 and 28. In sum, while there is a significant decrease in the share of siblings born within a very short time interval (less than 16 months apart from each), the increase of the share of siblings whose age gap is between 17 to 28 months is even larger leading to an overall reduction in spacing between births. Lalive and Zweimüller (2009) conclude from their findings that the 1990 reform not only affected the spacing between subsequent births (tempo or timing effect), but is likely to have actually increased completed fertility (quantum effect). Their results are robust to sensitivity and placebo checks (controlling for pre-birth characteristics; excluding births one week before and after the reform; running a placebo experiment using June 1987 and July 1987).

However, another study by use data on individual, parity specific birth records and who also investigate the effect of the 1990 parental leave extension on subsequent fertility behaviour, do not find any medium-term effects on total number of children within ten years after the reform. Nevertheless, this

latter study confirms the findings by Lalive and Zweimüller (2009) on changes in spacing between first and second births. It furthermore demonstrates that these changes apply to birth intervals between second and third births as well. More specifically,

of very short birth intervals of 15 to 20 months and increases the share of births occurring after 21 to 26 months following the last birth. The difference between the two studies lies mainly in the different types of data and samples: while Lalive and Zweimüller (2009) extract the relevant fertility information from administrative Social Security records and is thus limited to a certain subpopulation of women (ever employed in jobs subject to social insurance contributions; public sector workers as well as self-
(2009) involves information on all live births by women residing in Austria and is thus comprehensive. However, since the reform in 1990 could by definition only affect working women, it is likely that the latter analysis provides less precise estimates (since it is based on working (=eligible) and non-working (=non-eligible) women). On the other hand, information on birth events constructed from the Social Security database could be also imperfect, since births occurring before the first job are not captured.

One of the first studies to analyse the effect of the reform in 1990 on subsequent fertility was conducted by Hoem, Prskawetz and Neyer (2001a, 2001b). Using the Austrian Family and Fertility Survey from 1995/1996 to analyse trends and changes in *third births* in Austria from 1960 to 1996, they find that the July 1990 reform had a *tempo* effect on third births (at least in the short run), in the sense that the general trend towards postponement of third births was temporarily reversed. Since their analysis does not go beyond the year 1996 it is impossible to draw any conclusions regarding the persistence of this effect over time. Hoem, Prskawetz and Neyer (2001a, 2001b) attribute this narrowed spacing between second and third births to the implicit speed premium (automatic renewal period; waiver of employment requirements) of the new parental leave legislation.

3.2.3 Heterogeneity of effects across population subgroups

To test for differential reactions to the reform across subgroups, Lalive and Zweimüller (2009) run separate analyses for high and low wage earners (women earning above or below the median daily income one year before the birth) on the one hand and for white- and blue-collar occupations on the other hand. Both groups, high-

and low-wage mothers react to the extended leave period with higher fertility rates in the short term (within three years). However, mothers with low pre-birth wages react almost twice as strongly as mothers with high pre-birth wages: the probability of having an additional child within three years after giving birth increases by seven versus four percent respectively. Interestingly, a more detailed timing analysis reveals that high-wage mothers respond by significantly reducing their very short-term fertility (within 16 months after the previous birth), whereas this short-term postponement effect is much smaller for low-wage mothers. These group differences become even more pronounced when looking at the long-term fertility effect: while low-wage mothers have an excess fertility of five percent after ten years, there is no significant change in long-term fertility for high-wage mothers. Hence, high-wage mothers seem to have reacted to the extended leave period mainly by changing the spacing between births, while low-wage mothers have also increased the total number of births in the next ten years. Lalive and Zweimüller (2009) hypothesise that financial support during the parental leave period is supposedly more relevant to low income families by mitigating potential financial distress, whereas the job protection guarantee might be more relevant for career-oriented women with higher wages and more specialized skills and higher costs of job lost. In contrast to these stark differences in fertility behaviour, there seem to be no differential effects on short- and medium-term labour market outcomes apart from a slightly stronger short-run earnings reduction for low-wage women.

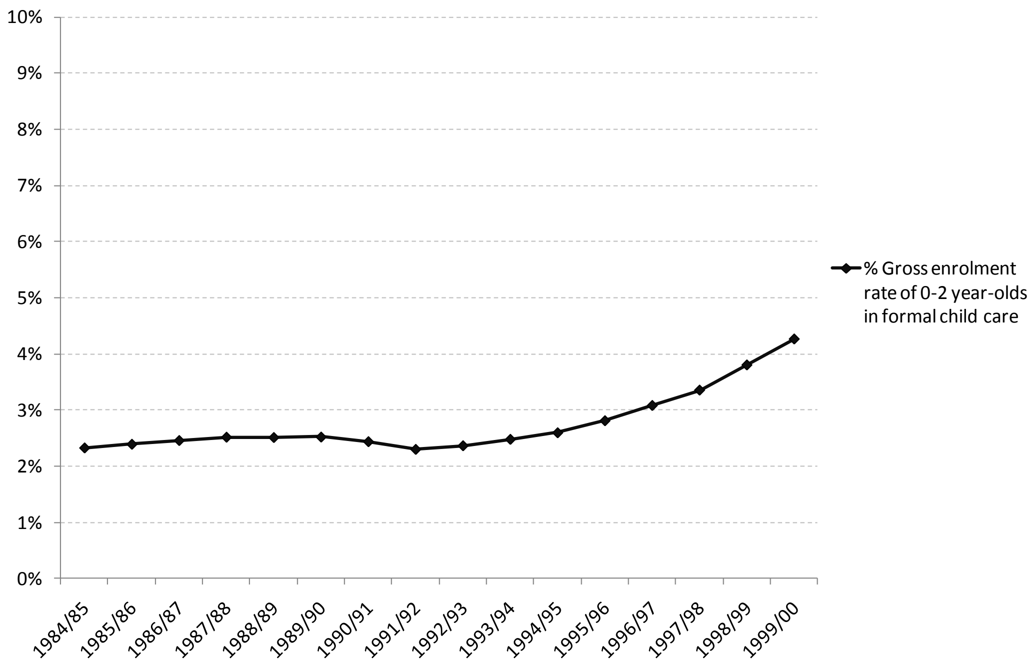
A different picture emerges when assessing differences between white- and blue-collar workers: both groups react quite similarly to the more generous leave regulations in terms of their short- and long-term fertility outcomes, even though white-collar mothers tend to adjust their timing more than blue-collar mothers. Under the more generous parental leave regime a larger fraction of white-collar women has their second child already within 28 months after the last birth (within the prolonged automatic renewal period of now 28 months). However, white-collar workers tend to react more strongly to the reform than blue-collar workers in terms of labour market outcomes. Especially in the short-run, treated white-collar workers have lower earnings and are less likely to return to work than white-collar control mothers. And although in the long-run there are generally no effects for either of the occupation groups, white-collar mothers have a slightly lower probability of having returned to work within ten years after giving birth (there are no differences in terms of long-term employment or earnings outcomes).

3.3 The availability and usage of formal and informal child care facilities

3.3.1 Child care for children from 0-2 and 3-6 years of age

The potential effect of maternal employment on child development depends crucially on the type and quality of care children are exposed to during the working hours of their mothers. Furthermore, the availability of centre-based or other forms of non-parental care (including care by relatives) for pre-school children generally facilitates the reconciliation of work and family life of parents and especially mothers and might therefore affect female labour force participation. This is why it is important to understand the extent of the availability and usage of child care in Austria around the time of the reform. Unfortunately, comprehensive child care information and in particular separate information by different types of child care formal and informal in Austria for the year 1990 or before is extremely scarce.

Figure 4: Gross enrolment rates of zero- to two-year-olds in formal child care (% of children in respective age group)

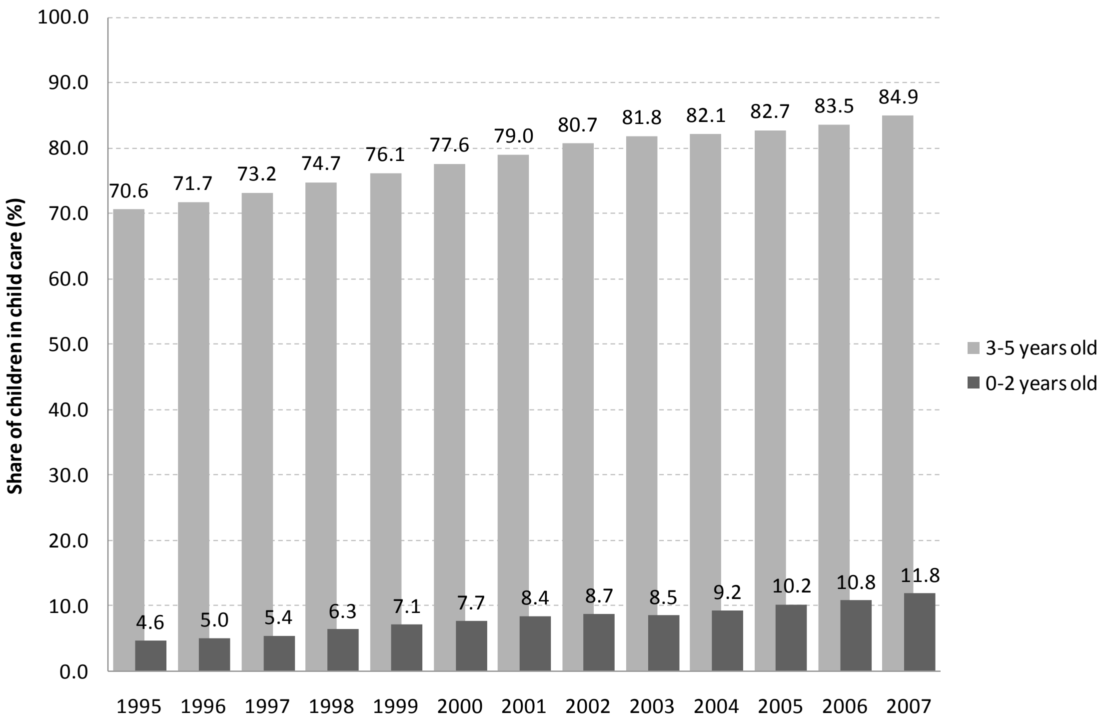


Notes: Gross enrolment rates of zero- to two-year-olds based on own calculations, dividing absolute numbers on children enrolled in formal child care by the sum of all births from the previous two years. Data sources: Statistik Austria (2010) and Human Fertility Database.

Figure 4 depicts the development of the enrolment rate of children younger than three years of age using an indicator that was calculated based on information on the number of births in the previous two years and enrolment numbers of zero- to two-year-olds in kindergarten (children enter formal care and school typically in September). According to this measure only about 2.5 percent of this age group participated in formal child care.[19]

Furthermore, there are official statistics on the provision of and participation in formal pre-school child care starting from 1995. While the numbers for zero- to two-year-olds are low, the enrolment rates of children aged three to six in formal childcare (Kindergarten) has been much higher and increased further over time (from around 70 percent in 1995 up to 85 percent in 2007, see Figure 5).[20]

Figure 5: Share of children in institutional day care over time (average for Austria)



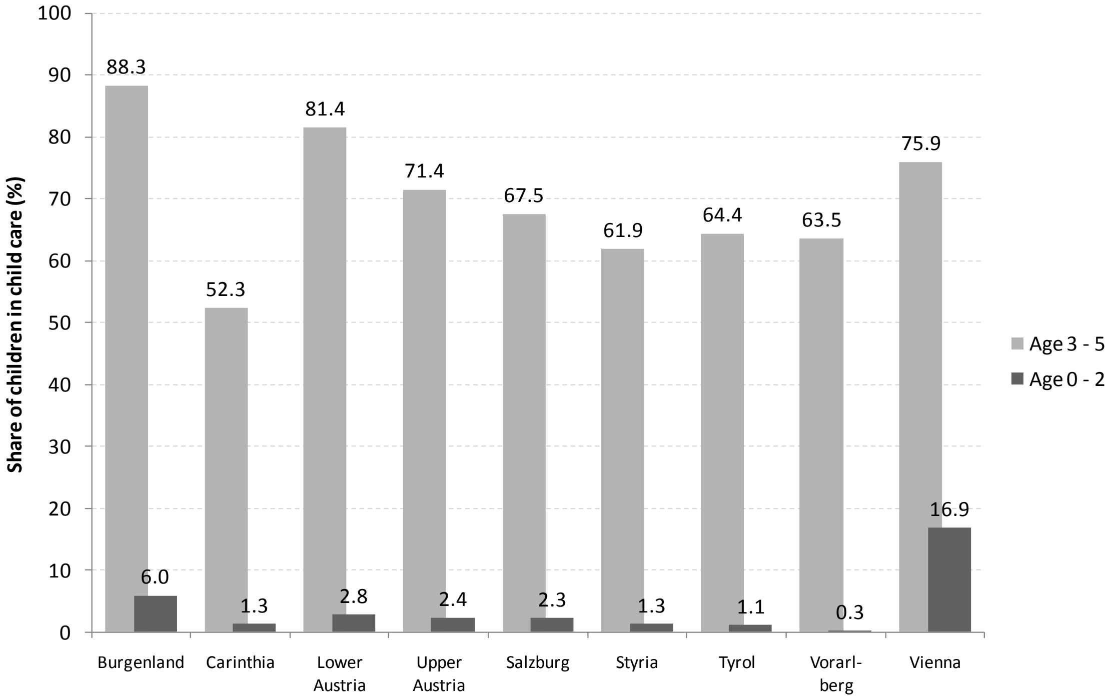
Source: Data from Statistik Austria (2010).

[19] These numbers tend to be biased slightly downwards, since the denominator does not take into account infant mortality. However, the positive trend in the figure is not driven by changes in infant mortality.
[20] Unfortunately these official statistics on enrolment in pre-school day care centers are only available starting from 1995.

Nevertheless, these rates are still below the 90 percent target rate set by the European Union for 2010 (Stanzel-Tischler and Breit 2009). The supply and usage of institutional child care for children under three years is even lower, but has traditionally been extremely low in Austria: in 1995 only 4.6 percent of children in this age group were registered in day care centres. Given the high Austrian female labour force participation rate it is not surprising that a survey in 1995 revealed excess demand for institutional day care for under three-year-olds (Dörfler 2004). Alternative enrolment figures of child care ratios stem from the micro census 1995 which collected information on overall external child care (institutional facilities as well as child minders or independent child groups). According to the micro census data, six percent of children under the age of three were in non-family day care, out of which about one third (36 percent) was provided through non-institutional child care arrangements (BSMUJ 1999a, p. 534). An OECD report (OECD 1989) provides numbers on enrolment rates in formal pre-primary education at age two in 1986/1987, i.e. before the parental leave reform: in Austria only 1.2 percent of the two-year-olds were enrolled in institutional care-centres, while in Norway and Finland the corresponding rates were 20.8 and 21.1 percent respectively (however, these numbers do not include other, less education oriented institutions like day mothers, crèches, day-care centres as well as informal care).

Apart from differences across age groups there is also a substantial variation in the proportion of children in institutional day care across regions in 1995 (see Figure 6). The region with the highest share of under three-year-olds in day care centres, namely 16.9 percent, is Vienna, the Austrian capital and the only city in the country with more than one million inhabitants (about 1.5 million according to the 1991 Census). The corresponding numbers of other regions of the country are much smaller and the day care shares ranged between 0.3 and 6.0 percent. Even though the overall ratio of children in formal care has increased over time this regional pattern has remained unchanged.

Figure 6: Share of children in institutional formal day care across regions in the year 1995



Source: Data from Statistik Austria (2010).

A special section on child care in the micro census from 2002 furthermore reveals that 9.8 percent of children under the age of three were enrolled in formal child care facilities (BSMG 2003). Within this group the highest fraction (40.4 percent) of children attended public pre-kindergarten child care (Kindergarten/Krippen), 21.9 percent attended private kindergarten, 23.7 percent were looked after by day mothers and 8.3 percent attended play or child groups. 70.6 percent of these children attended the child care facilities on five days a week, 29.3 percent on four days.

Unfortunately, information and data on use of informal (unregulated and not registered) child care provided by relatives, friends, neighbours, babysitters in Austria is very sparse. Nevertheless, this form of child care seems to play a non-negligible role in Austria, as 20 percent of children under the age of three are cared for by unpaid child-minders in a typical week in 2008 (OECD Family database 2010 (OECD 2010, Table PF3.3.A)).

Against this background of limited availability and low enrolment rates of zero- to two-year-olds in formal non-family day care in Austria even in the years 2002 to 2008, it seems unlikely that the reform in 1990 and the prolonged duration of the

parental leave implied a substitution from formal to maternal child care. Instead, the numbers tend to support the idea that the reform generated a substitution from individual informal day care arrangements (by relatives, friends, neighbours, etc.) to maternal care among those mothers who extended their leave period from one to two years as a reaction to the reform.

3.3.2 Time use of parents

An analysis of the interrelationship
participation and child care activities on the Austrian Time Use Survey from 1992 produces three interesting and relevant findings (Neuwirth 2004). First of all, and not surprisingly, increased market working time is associated with less time spent on child care activities. However, there is only an imperfect substitution between working and child care time: one additional hour of own labour market activity reduces maternal child care by only about 10 to 20 minutes. Furthermore, daily time devoted to child care rises with educational level ceteris paribus, i.e. holding market working time constant: mothers with university degrees spend on average 50 minutes per day more on child care activities than mothers with compulsory education. Time spent on child care is also higher for under three-year-olds than children aged four to six.

3.4 Possible effects of the Austrian reform on child outcomes (possible mechanisms)

Combining these identified reform effects with the theoretical considerations outlined in Section 2.1 there seem to be two potential channels through which the parental leave extension in Austria could have affected child development: first, a quality channel which works through the potentially superior quality of maternal care as opposed to alternative forms of child care, and second, the fertility channel , which works through changes in the fertility behaviour, i.e. the number of children and the spacing between births.

To summarize, the parental leave reform in 1990 significantly raised the time new mothers stayed at home after childbirth
birthday, but it did not affect medium- to long-run income and labour market outcomes of the average mother (thus, it is unlikely that the reform exerted a negative income effect on child outcomes driven by medium- to long-term income losses; if anything,

the leave extension generated a relatively small and only short-term income loss caused by foregone earnings during the additional leave months).[21] At the time of the reform the prevailing form of child care for children under two years of age was almost exclusively informal care provided by grandparents or other persons. Under the hypothesis that maternal care is superior to informal care for very young children, one would expect that the prolonged parental leave period had a positive impact on child outcomes. This could be especially true for better educated mothers if they are able to provide higher quality and more productive maternal care, for instance, through better access to knowledge on how to foster cognitive development of children (Grossman (2006) provides an overview and a critical discussion of the theoretical models as well as the empirical evidence regarding the effect of education on nonmarket outcomes, in which he distinguishes between *productive* and *allocative* efficiency, both of which are increasing in human capital and education).

However, the reform also altered the fertility behaviour of parents, most unambiguously the time interval between adjacent births: although the incidence of extremely short birth intervals declined, the average birth interval between first and second as well as second and third child decreased due to the incentives generated by the automatic renewal period. If shorter spacing between births reduces the time and material resources that are allocated to each child, this effect could have negative implications for child development and cognitive outcomes. On the contrary, if the relation between spacing and child outcomes is non-linear, a positive effect of the reduction in extremely short birth intervals could outweigh a negative effect from the average reduction in spacing. There is also some empirical evidence that the reform might have increased the total number of births per woman. If this is the case, this effect could work in the opposite direction and diminish or reverse a potentially positive time effect.[22]

---

[21] However, for those women who would *not*
the old regime anyways, the prolonged parental leave payments might have implied a *gain* in short term income as these mothers would not have earned *any* income during the second year. Furthermore, it is difficult to assess to what extent this short-term income loss from working mothers translates into a potentially negative income shock for the child: if non-parental child care is costly, a fraction of the
will be spent on child care and it is not clear whether the remaining amount of maternal income is actually larger or smaller than the parental leave payments. Only if the reform leads to a significant short-term income loss, there might be negative effects on child outcomes which could work against any positive time effect.

[22] The same could be true, if the reform exerts a significant negative income shock for families (see discussion in footnote 19).

Moreover, the reactions to the reform differed across population subgroups: high-wage mothers reacted less strongly to the extended leave period in terms of fertility behaviour than low-wage mothers (see Section 3.2.3). For the group of high-wage mothers there is no significant rise in overall fertility (no potentially negative quantity-quality trade-off), there is a much more pronounced decrease in extremely short birth intervals of 16 months (which should have a positive impact on child outcomes) and a weaker increase in short term fertility (within three years). In contrast, low-wage mothers have an excess fertility in the short and in the long-run by about seven and five percent respectively and there is no reduction in the incidence of extremely short birth intervals. These fertility effects are more likely to offset any positive time effect of the prolonged leave duration on child outcomes of low-wage mothers.

Overall, it seems that any positive effect from increased maternal time should be more easily detectable among children of high-wage mothers who altered their fertility behaviour only very modestly. For the children of low-wage mothers the effect of the parental leave reform is a mixture of counteracting changes of maternal care, increased family size and changed spacing between births.

4 Empirical approach

This section describes the empirical method used to investigate the link between the extended parental leave and cognitive child outcomes. It discusses the corresponding identifying assumptions as well as several refinements and potential problems with the approach.

4.1 Difference-in-Difference estimator and identifying assumptions

Given the unexpected and strict implementation of the prolonged parental leave period for all children born on July 1, 1990 or later, it is possible to use a Difference-in-Differences (DID) analysis to identify the effect of extended maternal care on child outcomes (like in the analysis of Germany by Dustmann and Schönberg 2010). The DID regression specifications are the following:

$$y_i = \alpha + \beta_1 Post\,June + \beta_2 y2006 + \beta_3 Post\,June * y2006 + \theta_m birth\,month + \varepsilon_i \qquad (4.2)$$

$$y_i = \alpha + \beta_1 Post\ June + \beta_2 y2006 + \beta_3 Post\ June * y2006 + \theta_m birth\ month + \mu X + \varepsilon_i \quad (4.3)$$

$\boldsymbol{y_i}$ is the measure for cognitive child outcome (i.e. test scores from standardized tests from the two years 2006 and 2003 covering the cohorts of children born in 1990 and 1987, respectively, see Section 5 for data description), *Post June* is a dummy variable taking on the value 1 for all children whose birthday is on or after July 1 (July December) and the coefficient $\boldsymbol{\beta_1}$ captures all possible permanent and general differences between children born in the first and the second half of a given year; *y2006* is a dummy variable for the year 2006 that controls for the common trend between the two test years 2006 and 2003; the interaction effect between *Post June* and *y2006* identifies all children whose mothers were affected by the reform and eligible to a longer parental leave $\boldsymbol{\beta_3}$ is the coefficient of interest and measures the treatment effect; to account for possible season of birth effects as well as age effects the regressions include a set of *birth month* dummy variables.[23] To control for possible differences in sample composition over time the more refined specification (4.3) contains additionally a set of parental and other background characteristics (*X*).

If the assignment into treated (=post reform; 24 months PL) and control group (=pre reform; 12 months PL) is as good as random , a simple representation of the estimated treatment effect $\boldsymbol{\widehat{\beta}_3}$ (estimated by OLS) is

$$\hat{\beta}_3 = (\bar{y}_{2006,post} - \bar{y}_{2006,pre}) - (\bar{y}_{2003,post} - \bar{y}_{2003,pre}). \quad (4.4)$$

This is the difference in average cognitive outcomes (test scores) of children born after versus before the reform (whose mothers were eligible to 24 versus 12 months of paid parental leave respectively) less the difference in outcomes of children born before and after July 1, 2003 who were not subject to the reform.

The advantage of this DID approach is that potentially confounding systematic differences between children born before and after July 1 which could otherwise exert a bias are differenced out: First, the test scores used in the analysis stem from tests that took place within a certain month (e.g. April) and children born in January 1990 will be about 12 months older at the time of the test than children born in December 1990. If age in itself has a positive effect on outcomes, any potentially positive effects of the

[23] Figure A 1 in the appendix reveals a seasonal pattern in the number of births (on average there are more births per day in July, August and September than in the rest of the year). However, these seasonal trends are constant over the years and should thus be accounted for in the DID approach.

reform will be downward biased, since post-reform children are always younger than pre-reform children. Second, there might be systematic season of birth effects affecting the composition of children and their parents over the year. If certain types of couples are more likely to have babies in particular months of the year this might also impact upon the distribution of test scores across birth months.

One of the most crucial identifying assumptions for this approach is that assignment into treatment was indeed as good as random , i.e. that mothers could not self-select into treatment or control group. This basically requires that mother could not manipulate the date of childbirth around the cut-off date of July 1, 1990. Lalive and Zweimüller (1999) provide several arguments and evidence in support of the assumption of random assignment: first, an assessment of newspaper reports about a potential reform of parental leave duration revealed that the public discussion did not start before November 11, 1989 and that it was not clear until April 5 whether and when such a reform would be implemented. This timing of policy decision and implementation makes anticipatory adjustments to fertility plans highly unlikely, especially when taking into account that successful conception and date of childbirths cannot be perfectly controlled and planned by parents. Furthermore, as also argued by Würtz Rasmussen (2010), it is biologically infeasible to postpone the date of delivery (it is easier to give birth at an earlier date). The exceptions are births by Caesarean sections. However, an analysis of number of births during the days shortly before and after the reform did not indicate a higher density of births on July 1 or the days after (Lalive and Zweimüller 2009). Another assumption is the *common trend* assumption which requires that seasonal patterns or age effects are constant across years.

However, the common trend assumption might be problematic if there are changes over time. Certainly, this assumption becomes less restrictive if one limits the sample to children born extremely close to the cut-off date as these children are very similar in age as well as in season of birth (like in a Regression Discontinuity analysis)[24]. Another advantage of narrowing the window of birth months before and after the reform would be that these children are more likely to face identical kindergarten and schooling regulations and rules, e.g., the Austrian school year typically runs from September to August (see also Section 4.3). Furthermore, their mothers were

[24] This relates to the general idea behind a Regression Discontinuity Approach which relies on the similarity of composition of treatment and control group in terms of observed and unobserved characteristics who only differ with respect to the assignment into treatment.

exposed to similar macroeconomic conditions and labour market developments (this argument in favour for a narrow sample around the cut-off date is given by Lalive et al. 2010). However, this strategy would require a very large data set. Unfortunately, the available test scores data base does not satisfy this condition.

Hence, given the data at hand, there is a *trade-off* between limiting the analysis on children who are as similar as possible (which would also reduce the likelihood of violating the common trend assumption) and having a sufficiently large sample size. Therefore, each specification will be estimated several times while successively narrowing down the window of birth months (starting from the birth months March to October and narrowing it down to May to August).

Since the data neither allow identifying children whose mothers were actually eligible for the more generous parental leave entitlements nor contain information on actual duration of leave taking of mothers, the estimated effect will represent the intention-to-treat effect of the reform (i.e. the reduced form effect of being eligible to 24 instead of 12 months of parental leave; see explanation and discussion by Baker and Milligan (2008, 2010) and Dustmann and Schönberg (2010)). More specifically, the estimated intention-to-treat effect is the average causal effect of the *assignment* of treatment on the outcome. The estimation procedure will simply compare outcomes of students that were randomly assigned to the treatment group or the control group (i.e., 24 or 12 months of parental leave), without accounting for the fact that neither all children in the treatment group actually received the treatment nor all children in the control group did not receive the treatment of prolonged parental leave. This intention-to-treat effect will be a lower bound estimate of the effect of prolonged parental leave and maternal care on child outcomes (compliers: mothers who adjust the length of the parental leave in response to the extended entitlements), since it is estimated on the full sample *including* children of mothers who did not change their behaviour because of the reform (non-complying mothers; see Angrist, Imbens and Rubin 1996). This latter group consists of always-takers (mothers always staying at home irrespective of the parental leave regime), i.e. non-working mothers as well as working mothers, who stop working post-birth for much longer than the granted parental leave period irrespective of the actual legislation, and of never-takers (mothers who return to work very early irrespective of the generosity of the system), i.e. non-eligible working mothers, e.g. self-employed, or working mothers would return to work immediately after the compulsory

maternity leave period independent of the actual leave provision. Furthermore, the intention-to-treat estimate could be a combination of different channels through which the reform might have affected child outcomes (see discussion of channels in Section 3.4). If information on *actual* maternal leave taking duration had been available in the data, the reform cut-off date could have been used as an instrumental variable in order to estimate the local average treatment effect of the reform.

4.2 Refinement and subgroup analysis

One way to get closer to the actual effect of the reform would be to restrict the sample to children whose mothers were actually eligible for parental leave and hence affected by the reform. Although direct information on the working status of the mother is not available in the data, it is possible to infer average labour force participation rates from educational attainments of mothers by splitting the sample into sub-groups of mothers with comparatively high or low labour force participation rates and hence a higher and a lower likelihood of parental leave eligibility (a similar approach is adopted in Baker and Milligan 2008).

Table 3 provides information on employment rates of mothers aged 18 to 39 with a baby younger than one year. The ratios were calculated using data from the Austrian Census 1991 (May 15, 1991). The employment measure includes women who are currently absent from work due to maternity or parental leave in the numerator (ratio of employed women to all women in the respective age group). Note that the Census 1991 was conducted before the first cohort of post-reform mothers started their second year of parental leave (by July 1991).

According to these figures women having completed post-secondary or tertiary education (having employment ratios of above 80 percent) will be categorized as High labour force participation group (High LFP) , and those with lower ratios will be subsumed as Low labour force participation group (Low FLP) . The latter group comprises women whose highest educational degree is higher secondary school or less. On average, these groups have an employment ratio of about 84 percent (High LFP group) and 70 percent (Low LFP group).[25] When including women born abroad in the calculations the numbers become slightly lower, but the overall ranking remains the

[25] When including unemployed in the economically active become 89 and 78 percent, respectively.

same (on average 81 (67) percent for the High (Low) LFP group; this might be related to lower labour force participation rates or higher unemployment rates among foreign-born individuals). Unfortunately the census data does not allow a distinction between employed and self-employed or unpaid workers, and thus these numbers probably slightly exaggerate the share of mothers with employment contracts granting parental leave eligibility (possibly especially so for the less educated).

Table 3: Female employment ratio by highest education completed, Census Austria 1991

		Born in Austria		Born anywhere	
ISCED	**Highest education completed**	**Employment ratio**	**N**	**Employment ratio**	**N**
1. Group of mothers with lower labour force participation					
2	Compulsory secondary school	60.8%	2,020	57.2%	2,528
3C	Intermediate technical & vocational secondary school (short form)	67.7%	136	67.2%	137
3B	Upper secondary	75.3%	3,828	74.6%	3,970
3A	Higher general secondary school	65.9%	437	60.8%	523
	Total	*70.0%*	*6,421*	*67.3%*	*7,158*
2. Group of mothers with higher labour force participation					
4	Post-secondary (not tertiary) (Intermediate or higher technical & vocational secondary school)	81.7%	699	79.7%	744
5B	Post-secondary college (tertiary)	90.7%	333	90.0%	341
5A/6	University, Polytechnic (tertiary)	81.6%	305	76.8%	358
	Total	*83.9%*	*1,337*	*81.4%*	*1,443*

Notes: Subsample of all mothers aged 18 to 39 years with a child younger than one year (Austria, Census date May 15, 1991). The employment ratio is calculated as the ratio of persons working for an employer, self-employed persons, unpaid workers engaged in the production of economic goods, and persons who have a job but were temporarily absent for some reason (e.g. maternity or parental leave) divided by the total number of people in this age group. Employment measure does not include unemployed individuals, since the focus of the following analysis is on working mothers. Including unemployed women in the employment measure changes the ratios only slightly (they become larger). The educational classification is according to ISCED 1997. Source: Census Austria 1991 downloaded from the Integrated Public Use Microdata Series International (Minnesota Population Center 2011); own calculations.

A division between these two groups of mothers according to educational attainment will not only differentiate between supposable eligibility statuses (women with higher education being more likely to be affected by the reform because of higher employment rates). In addition, if more educated mothers are more career-oriented and have stronger incentives to return to work immediately after the end of the granted period of job protection their return-to-work decisions might follow the official rules more strictly (the period of parental leave entitlements might be more binding for women with a stronger attachment to the labour force). If this is the case, an estimation based on children from this subgroup of mothers will reflect the extension of the specific effect more clearly (mothers with lower attachment to the labour force might stay away from work for a longer period irrespective of the granted parental leave period). Moreover, a higher level of education might have a stronger and positive impact on child development in itself, if mothers with higher education are able to provide better quality time and care at home than less educated mothers (or informal child care).

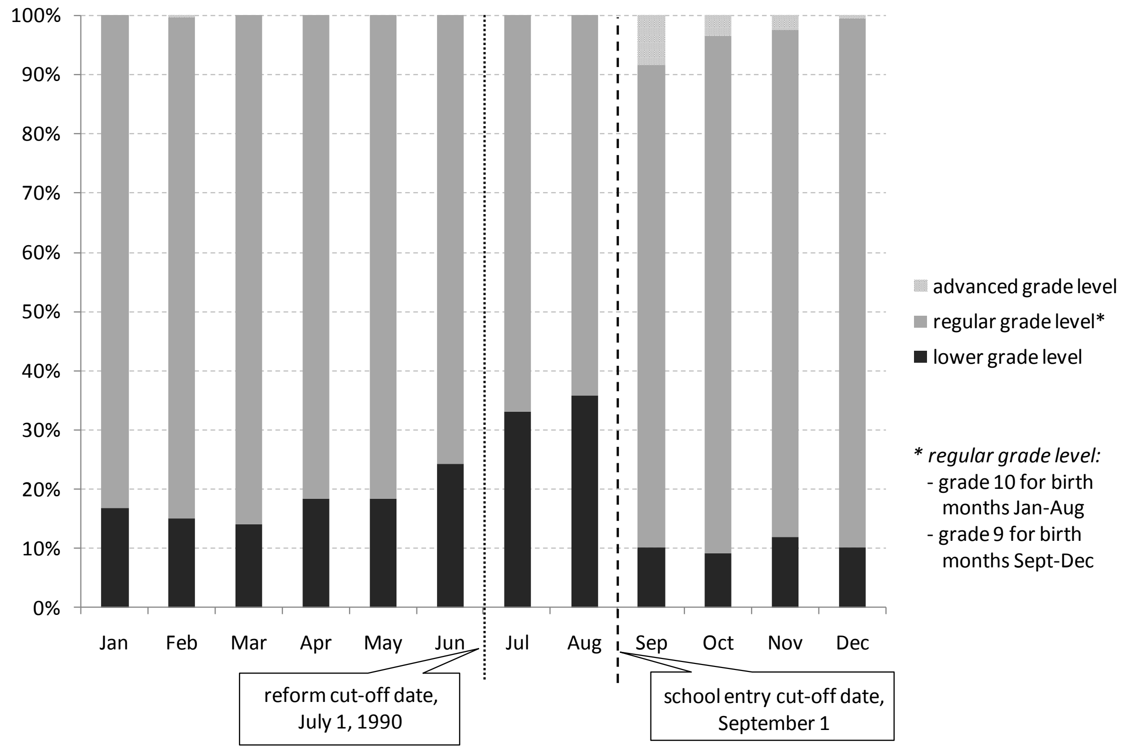
4.3 Potentially confounding effects: School-entry cut-off age

One additional challenge for the empirical analysis is the proximity of the reform cut-off date July 1, 1990 to the Austrian school entry cut-off date which is the 1st of September of each year. More specifically, all children having their sixth birthday before September 1 of each year are obliged to start school in the same year (school starts typically at the beginning of September) unless they are officially considered as not yet ready for school given their cognitive skills (level of maturity according to professional opinion). However, until 1999, enforcement of this school entry cut-off date was not strict, conceding some discretionary power to *parents* regarding the decision whether their child was ready for school or whether it should start one year later. Overall, the fraction of children enrolling late tends to be particularly high in the birth months immediately preceding the cut-off date. Furthermore, children born in July and August who enter school at the regular age will always be the youngest in their class. Simple age effects (biological, cognitive development) as well as more complex psychological peer effects might increase the likelihood that these children perform relatively worse than others and might thus be at higher risk of repeating a grade (recent empirical accounts of such relative age effects within classrooms are given by the

following studies: Bedard and Dhuey (2006), Black, Devereux and Salvanes (2011), Mühlenweg (2010), Schneeweis and Zweimüller (2009) and Sprietsma (2010)).[26]

These two effects late enrolment and grade repetition lead to a very unequal distribution of the share of children enrolled in their *regular* grade level given their birth date across all birth months. Figure 7 plots the fractions of 15 year old children in Austria (for the reform cohort 1990) who were enrolled in a regular, advanced or lower grade level by month of birth. The categorization into regular, advanced and lower level is based on a comparison of their actual grade level with the grade level they should be in according to their birth date (based on the PISA 2006 data; for data description see Section 5.1). More specifically, in this particular case all children born before September should be enrolled in grade 10, while all children born between September and December should be enrolled in grade 9.

Figure 7: Grade progression by birth month



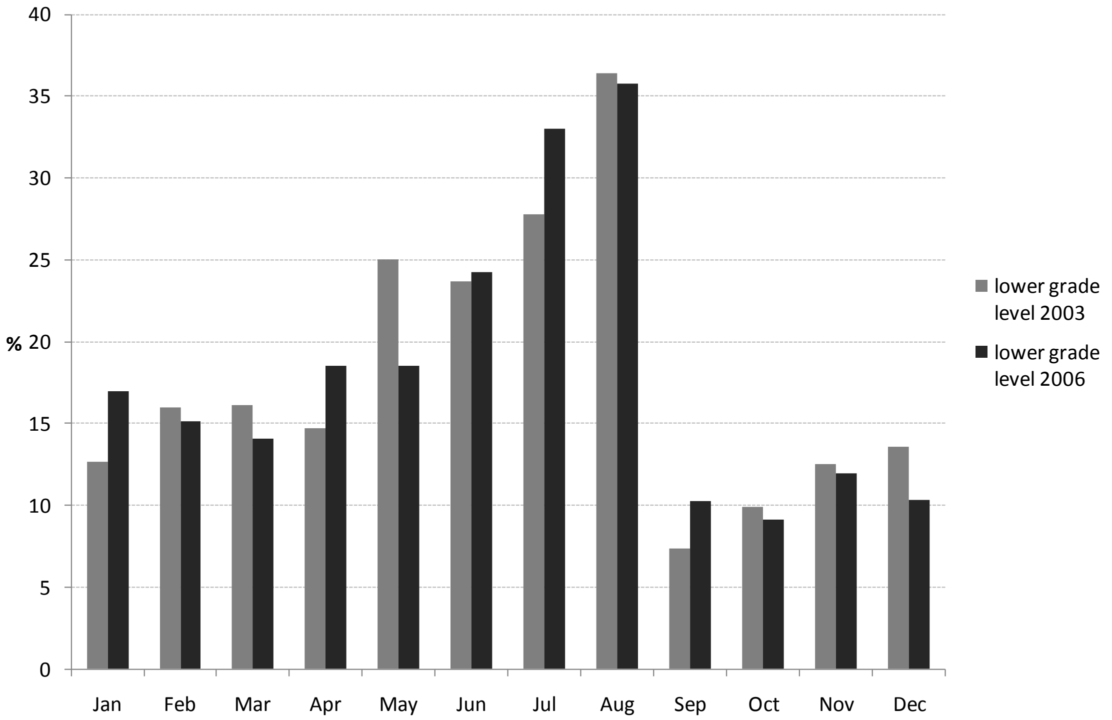
Source: PISA 2006; own calculations based on full sample included in the analysis and using student weights calculated by the data providers (N= 4,372; see sample description in Section 5.2).

[26] The cited literature also discusses other potential relative age effect channels. On the one hand, relatively younger children might benefit from the fact that they cover certain material at a younger age than their older classmates. On the other hand, there could be a negative effect if younger children imitate risky behaviour from their older peers.

The figure reveals that the fraction of children in lower than regular grade levels exceeds 30 percent in the months July and August (33 and 36 percent respectively), whereas in June and May the fraction drops to 24 and 19 percent. Under the plausible assumption that children in lower grade levels will perform comparatively worse in the PISA tests than children in higher grade levels since they have completed fewer years of formal education and covered less material the *average* test scores of children born in August and July will be lower than for children born earlier in the year. The same is especially true for children born between September and December 1990: among these children about 80 percent are in grade level 9 and hence their average test scores will be lower than those in the preceding birth months.

The comparison of the distribution of children in lower than regular grades across birth months for the cohorts 1990 and 1987 (PISA data 2006 and 2003) in Figure 8 reveals that this phenomenon is relatively consistent over time. This is an important prerequisite for the DID estimation strategy.

Figure 8: Distribution of students in lower than regular grade level (cohorts 1990 and 1987)



Source: PISA 2006 and 2003; own calculations based on full sample included in the analysis and using student weights calculated by the data providers.

The compositional changes in grade level attendance across birth months and especially to the right of the reform cut-off date seems to call for a limitation of the analysis to the birth months June and July only. Unfortunately though, as already mentioned earlier, this strategy is not possible due to the small sample size. To test to what extent the empirical results are sensitive to this school entry cut-off date, as a further robustness check the following analysis is also run only including children born before September (but still gradually extending the sample to the left of the cut-off).[27]

An alternative sensitivity test will restrict the sample to children who are in the regular grade only. However, since current grade level of a student might be endogenous to his cognitive skills, these results can only provide suggestive evidence and have to interpreted with caution.

5 Data on cognitive outcomes of the 1990 birth cohort in Austria

An analysis of the effects of the 1990 parental leave extension on medium-term cognitive child outcomes requires micro data on scholastic achievements from students born shortly before and after the reform in 1990. However, nationally representative

par

1990 and 1987 respectively (see motivation for and advantages of the DID approach in Section 4.1).

There are several important features of the PISA data which make it especially suitable for the analysis: first of all, the PISA data provides results from *standardized tests* of cognitive skills in terms of reading, mathematical and scientific literacy. The focus of PISA is less a pure assessment of curriculum based knowledge, but more an evaluation of general skills needed for adult life and of the ability to apply knowledge to real-life problems. Second, the tests are administered to a nationally representative sample of 15-year old students independent of their current grade level in school. In contrast, other international studies like TIMSS and PIRLS assess students in particular grade levels, e.g. 4th and 8th grade, and are thus not representative for a particular birth cohort. Comparisons of outcomes across birth months would be biased if, for example, the propensity of grade retention or early or late school entry differs between children born closer or further away from the school entry cut-off date (see discussion in Section 4.3). Third, the PISA data files contain important student-reported background information on the student (e.g., gender, birth year and month, nationality, attitudes), the , occupational information) and the school (e.g., school programme, location, school size, resources).

However, it should be mentioned that the Austrian PISA data has also several disadvantages: although the overall sample size in 2006 includes about 4,927 students, the relevant sample when comparing children born in particular months of the year becomes quite small (there are only about 350 children per birth month). Furthermore, there is no retrospective information on maternal labour market participation at the time of birth which prevents a clear identification of mothers who were truly eligible for parental leave. Moreover, information on exact birth dates is not available in the data due to strict data security and protection rules. The lack of information on exact birth dates in conjunction with the rather small sample size prevents any refinement of the analysis beyond the month level (e.g. children born shortly before or after the cut-off date cannot be excluded to test for robustness to potential sorting across the cut-off date).

The Programme for International Student Assessment was initiated and is coordinated by the Organisation for Economic Co-operation and Development

(OECD)[28]. The international micro data files are publicly available free of charge from the OECD web-site which also offers detailed information on the data, variables and survey methodology (see http://www.pisa.oecd.org/). In Austria, the PISA 2003 and 2006 tests were managed and carried out by the Federal Institute for Education Research and Innovation & Development of the Austrian School System (Bifie Bundestinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens).[29]

The students participating in the Austrian PISA test were sampled according to a two-stage stratified sampling design: the first stage involved a random selection of schools from the universe of all schools having 15 year old students (probability-proportional-to-size sampling, where size referred to number of 15-year-olds within school; the stratification also accounted for different school types and programmes); in the second stage a maximum of 35 15-year-olds were selected within each school to participate in the test (Breit and Schreiner 2007). The test itself was designed as a two hours paper-and-pencil assessment including multiple-choice and open answer questions plus a 30 minutes background questionnaire (OECD 2009a).

5.2 Sample selection and variables

Several observations were dropped from the original data in order to increase the cohesiveness of the data for the analysis. First, students whose mothers are highly unlikely to have been eligible for parental leave or affected by the reform are excluded: this relates to children who were not born in Austria and whose mother was thus unlikely to work in Austria at the time of the reform in 1990. Furthermore, a few observations had missing information on maternal education (parental education is reported by the student). These observations were dropped for two reasons: first, missing answers on maternal education could indicate absence of mothers due to death, separation, etc. which is problematic given the aim to evaluate the importance of *maternal* employment or care on child outcomes. Second, the level of maternal education is a crucial variable according to which the sample will be divided into two groups in the subsequent analysis and missing information on maternal educational

[28] Up to date, PISA tests were carried out in the years 2000, 2003, 2006 and 2009 and the number of countries participating in the test has increased from 43 countries (OECD and non-OECD) in 2000 to 65 countries in 2009.
[29] Further details on the implementation of PISA in Austria are available at http://www.bifie.at/pisa.

attainment prohibits the allocation of the student to a particular group. A few students *in schools for children with special needs* were also excluded from the sample to increase consistency across students and across years, since these children have a completely different curriculum and these children were administered special test items (Schreiner, Breit, Schwantner and Grafendorfer 2007).

The variables that are used in the analysis are described in Table 4. The main outcome variables measuring cognitive skills are test scores in mathematics, reading and science.[30] The test scores in each subject are rescaled by the OECD so that the mean across all participating countries is 500 points and the standard deviation is 100 points.[31]

Two other potential outcome variables are indicator variables on whether a student is in a lower than the regular grade level given their birth month and whether the student is enrolled in a school track which gives access to university or college education (academic track). In Austria, students are allocated to different educational tracks after the fourth grade (i.e. at age 10 or 11), which is relatively early compared to other European countries (Schneeweis and Zweimüller 2009). Starting with grade level nine, a further differentiation into specific tracks takes place; the PISA sample thus covers several different school types representing high and low track schools, which are also associated with different levels of PISA test outcomes. A list of these national school programmes and how these are categorized into academic track schools can be found at the bottom of Table 4. Again, since these different school programmes are associated with different achievement levels and since they have different orientations (vocational training versus academic education) enrolment into a particular school types is an endogenous outcome.

[30] Strictly speaking, the PISA data does not contain standard test scores, but so-
(five plausible values per subject) which are estimates of the overall proficiency of a student given his/her particular objective test outcomes, drawn from an underlying latent ability distribution of the student. As recommended in the PISA Data Analysis Manual (OECD 2009b), the following analysis is based on one of these plausible values (plausible value 1) for each subject (as has been done in another study by Schneeweis and Winter-Ebmer 2007). Also in line with the OECD recommendations all presented results are weighted using the provided student level weights. Standard errors are clustered by school programme, school location and gender to account for within group correlation of outcomes.

[31] More specifically, to ease comparison of results across survey years, the test results were rescaled such that the reading and mathematics reporting scales of 2006 are equal to those in 2003. The test results in science are rescaled such that the mean is 500 and the standard deviation is 100 for the 30 OECD countries that participated in PISA 2006 (see OECD 2009a, PISA 2006 Technical Report, pp. 157-158). However, any general differences in test scores across years and subjects will be accounted for in the analysis by including controls for years and by running separate regressions for each subject.

Table 4: Description of variables

Variable Name	Variable Description
School outcomes	
Mathematics score	Mathematics test score (plausible value 1)
Reading score	Reading test score (plausible value 1)
Science score	Science test score (plausible value 1)
Retained	= 1, if observed grade level is below grade level student should be according to this birth month (regular grade level is 10 for children born before September, and 9 for children born in September or later) (= 0, otherwise)
Academic track	= 1, if enrolled in school which enables student to attend university after graduation (= 0, otherwise)
Male	= 1, if gender is male; = 0, if gender is female
Age in years	Age in years at the time of the test
School location (dummy variables)	
City	= 1, if school is located in city , i.e. 100,000 to 1,000,000 inhabitants (= 0, otherwise)
Metropolitan area	= 1, if school is located in metropolitan area , more than 1,000,000 inhabitants (= 0, otherwise)
Mother's and father's educational attainment (dummy variables):	
Lower secondary	Lower secondary or less [→ low LFP group]
Upper secondary	Upper secondary [→ low LFP group]
Tertiary	Tertiary or Post-Secondary [→ high LFP group]
Educ. father: miss.	Educational attainment of father is unknown or missing
Migration background of family (dummy variables)	
Family type 2	= 1, if family speaks German at home, but at least one of the parents was born abroad (= 0, otherwise)
Family type 3	= 1, if both parents were born in Austria and home language is German (= 0, otherwise)
National school programme (Austria) (dummy variables):	
Vocational (low track)	Vocational (low track) (Hauptschule, Polytechnische Schule)
Apprenticeship	Apprenticeship training (Berufsschule)
BMS	Medium vocational school (Berufsbildende Mittlere Schule)
BHS	Higher vocational school (Berufsbildende Höhere Schule) [→ academic track]
AHS	Gymnasium, Higher general school (Allgemeinbildende Höhere Schule) [→ academic track]

implicitly accounted for by controlling for birth month), dummy variables indicating whether the school is located in an urban area (i.e. metropolitan area or large city; the base category are smaller towns with less than 100,000 inhabitants and villages),

variables for migration status of the family (whether the family speaks German at home and whether mother and father were born in Austria; the base category are non-German

speaking families). All these variables have been selected to be included in the analysis since they are comparatively unlikely to be endogenous to child characteristics (i.e., these variables were chosen as to present pre-birth characteristics of the family). This endogeneity problem is the reason why no information on current employment or current occupation of the mother is incorporated in the analysis (*current* refers to date of the respective PISA test).

5.3 Descriptive statistics and first raw comparisons

The mean values of the outcome variables and control variables of the PISA 2006 and 2003 data are presented in Table 5, separately for children born in May and June (before the reform in 1990) and for children born in July and August (after the reform in 1990).

As regards the other demographic characteristics, school loc

differences between pre and post reform children in 2006, which supports the important assumption that the reform was unexpected and that there was not systematic self-selection of particular types of parents or families around the cut-off date. In 2003 there seems to be a higher density of mothers with tertiary education among children born in July or August than born in May or June.

Table 5: Mean comparisons of outcomes and characteristics of students born in May/June versus July/August in the reform year 1990 and the control year 1987

	PISA 2006 (birth cohort 1990)				PISA 2003 (birth cohort 1987)			
	Pre reform	Post reform			Pre reform	Pre reform		
Birth months	**May-June**	**July-Aug.**			**May-June**	**July-Aug.**		
	Mean (1)	Mean (2)	Diff. (2)-(1)	Std. error	Mean (1)	Mean (2)	Diff. (2)-(1)	Std. error
Mathematics score *(Std. deviation: 92.8)*	521.20	519.80	-1.45	5.23	515.90	515.10	-0.74	5.92
Reading score *(Std. deviation: 101.1)*	507.20	500.70	-6.46	5.74	501.10	501.70	0.62	5.73
Science score *(Std. deviation: 90.6)*	526.90	523.70	-3.18	5.09	504.40	502.50	-1.87	5.65
Retained	0.21	0.34	0.13**	0.03	0.25	0.33	0.08**	0.03
Academic track	0.54	0.52	-0.02	0.03	0.46	0.51	0.04	0.03
Male	0.52	0.50	-0.02	0.03	0.50	0.49	-0.01	0.03
Age in years	15.92	15.75	-0.17**	0.00	15.93	15.76	-0.17**	0.00
City	0.41	0.41	-0.01	0.03	0.34	0.35	0.01	0.03
Metropolitan area	0.14	0.13	-0.01	0.02	0.16	0.14	-0.02	0.02
Lower secondary	0.09	0.09	0.00	0.02	0.13	0.09	-0.04**	0.02
Upper secondary	0.54	0.56	0.02	0.03	0.59	0.56	-0.03	0.03
Tertiary	0.37	0.35	-0.02	0.03	0.28	0.34	0.07**	0.03
Lower secondary	0.06	0.06	0.00	0.01	0.09	0.07	-0.02	0.02
Upper secondary	0.47	0.46	-0.01	0.03	0.51	0.48	-0.03	0.03
Tertiary	0.45	0.46	0.02	0.03	0.37	0.41	0.05	0.03
Educ. father: miss.	0.02	0.02	0.00	0.01	0.03	0.04	0.01	0.01
Family type 2	0.09	0.07	-0.02	0.02	0.06	0.06	0.00	0.01
Family type 3	0.85	0.86	0.01	0.02	0.90	0.91	0.01	0.02
School programme:								
Vocational (low track)	0.05	0.07	0.02	0.02	0.05	0.07	0.02	0.02
Apprenticeship	0.26	0.23	-0.04	0.02	0.26	0.24	-0.02	0.03
BMS	0.13	0.16	0.03*	0.02	0.20	0.17	-0.03	0.02
BHS	0.33	0.30	-0.03	0.03	0.31	0.32	0.01	0.03
AHS	0.23	0.24	0.01	0.02	0.18	0.21	0.03	0.02
Number of observations	*716*	*764*			*680*	*680*		

Notes: Estimations weighted by individual inverse probability weights provided in the PISA data set. The standard deviations refer to the 2006 data using the pooled sample from May to August. Significance levels are denoted by: ** p<0.05, * p<0.1.

6 Empirical results of the effect of the reform on cognitive skills

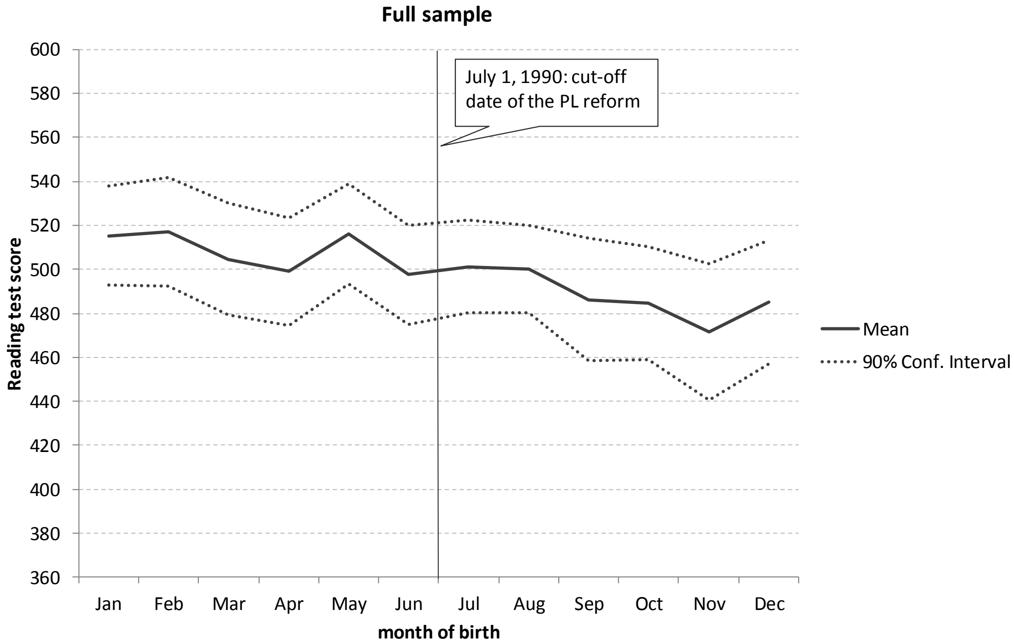
After showing some simple graphs on the distribution of the average test scores across birth months, this section presents and discusses the results from the Difference-in-Differences estimations.

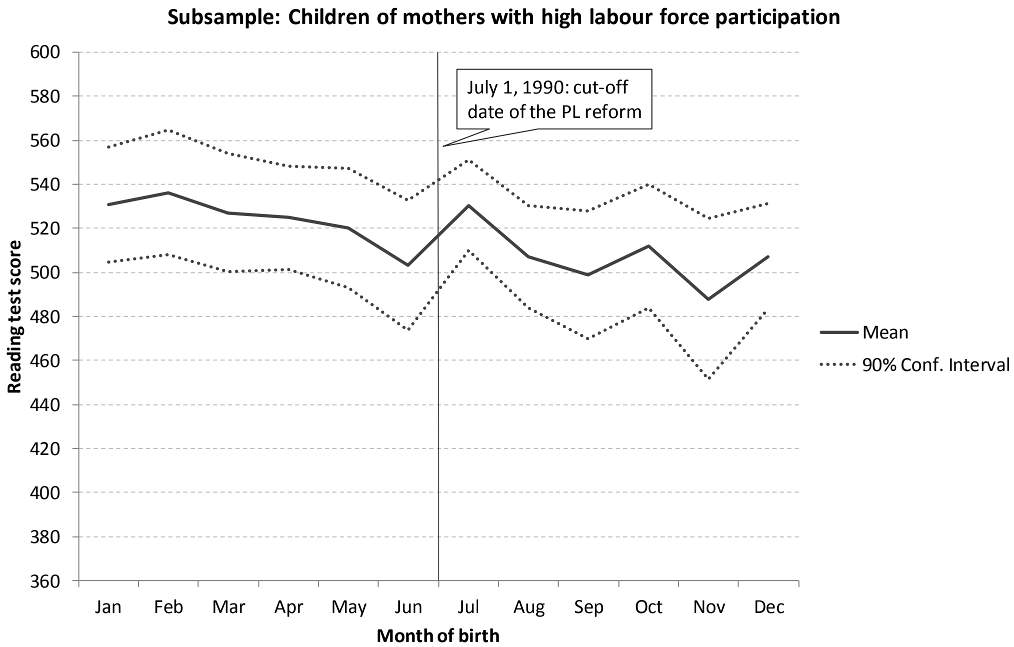
6.1 Graphical Results

Figure 9 helps to get a first impression and better understanding of the distribution of test scores across birth months. The graphical analysis is a simple tool to check whether there is a jump in test scores for children born after the parental leave extension (using reading test scores for this example). It seems that there is a general negative trend in average reading test scores across birth months that would be in line with possible age effects as discussed earlier. However, comparing the test outcomes immediately before and after the reform (i.e., June and July children) the absence of any marked differences suggests no positive reform effect. Nevertheless, this flat passage could imply a positive effect if it is the result of a positive reform effect neutralizing a negative age effect. Furthermore, the graph illustrates that children born between September and December 1990 perform on average worse in the tests. This is related to the fact that these children are in lower grade levels due to the school-entry cut-off regulations (children born in September or later enter school one year later).

In contrast, the picture looks different when the sample is limited to those children whose mothers were more likely to be affected by the reform due to a higher labour force participation rate (Figure 9, lower graph). Although the overall negative age effect remains, test scores of July children are much higher than for June or even May children and it seems as if the distribution was slightly upward shifted for post-reform children. Average test results of August children are lower than those for July children which would be in line with a negative age and school-entry effect (the share of August children who are in grade level nine instead of ten is higher than in July).

Figure 9: Average reading test scores across birth months (full and subsample)





Notes: The graphs plot the estimated birth month coefficients of a regression of test scores on a set of birth month dummy variables without additional controls (weighted using the individual students weights provided in the PISA data; the base month excluded from the regression is June) based on the full sample (upper graph) and on the subsample of children with mothers with higher labour force participation likelihood (lower graph). The 90 percent-confidence interval is based on estimated standard errors that were clustered by school programme, school location, and gender.

To test whether the test score gap between June and July children is significantly different from zero, a simple estimation was carried out based on the PISA 2006 before proceeding to the DID estimations. Test scores from the three different subjects were regressed on a set of birth months with June as the excluded base category, and other control variables on parental background. The estimated born in July coefficients are reported in Table 6 for separate regressions by subject, maternal labour force participation and child gender. The first two columns in Table 6 show the results from the full sample of mothers. As the previous graphs suggested the average test scores do not differ significantly between pre- and post-reform children in the full sample and this applies for all three tested subjects. Furthermore, the estimated coefficients change only marginally after controls for parental background are included in the regression (although standard errors become slightly smaller). Comparing the results by gender (columns 1 and 2, middle and bottom panel) reveals that there might be gender differences as the estimated coefficients for boys tend to be positive (but insignificant), while the estimated coefficients for girls tend to be negative. Restricting the sample to the subsample of mothers who were more likely affected by the reform (columns 3 and 4), confirms the graphical impression of a positive effect: the estimated effects become much larger and are significantly different from zero (at the five percent level) for reading and science. Again, the effects for boys seem to be larger than for girls. The separate analyses show that the differences remain significantly different from zero only for the male subsample. For completeness, the two columns on the right show the respective results for children from mothers with lower labour force participation rates. Most of the coefficients have a negative sign implying a decline in test scores after the reform; however, none of these effects is significantly different from zero.

Given the previous discussion of the potentially confounding age and school-entry effects, it is important to refine the analysis and to perform sensitivity checks in order to be able to draw more well-grounded and meaningful conclusions. This will be done in the next sections.

Table 6: Simple OLS regressions on the differences between June and July children using only data from the PISA test 2006 (including birth months May-August)

	(1) Full sample	(2) Full sample	(4) High LFP mother	(5) High LFP mother	(7) Low LFP mother	(8) Low LFP mother
			MALES + FEMALES			
Mathematics	3.043	3.289	15.687	15.276	-3.377	-2.920
	(7.002)	(6.208)	(11.040)	(11.074)	(8.358)	(7.519)
Reading	3.308	3.987	28.946**	28.559**	-10.111	-8.923
	(6.936)	(6.022)	(13.485)	(13.730)	(8.677)	(7.396)
Science	3.850	3.832	23.700**	22.113**	-6.424	-5.802
	(6.315)	(5.428)	(10.398)	(10.223)	(8.082)	(6.894)
Observations	1,480	1,480	523	523	957	957
			MALES			
Mathematics	11.734	8.615	27.893	21.853	2.882	1.934
	(8.957)	(8.014)	(17.287)	(14.846)	(8.820)	(9.099)
Reading	12.491	10.470	50.349**	44.261*	-7.795	-5.048
	(9.167)	(7.871)	(22.448)	(21.999)	(10.809)	(9.806)
Science	7.720	4.677	40.772**	34.568**	-9.886	-10.464
	(8.559)	(7.312)	(15.513)	(14.392)	(8.722)	(7.863)
Observations	752	752	265	265	487	487
			FEMALES			
Mathematics	-5.771	-2.289	3.492	5.848	-9.627	-4.493
	(10.041)	(9.823)	(14.652)	(15.520)	(14.027)	(13.150)
Reading	-6.051	-2.955	8.174	11.779	-12.219	-8.156
	(9.755)	(8.702)	(15.309)	(15.431)	(13.713)	(11.104)
Science	-0.118	2.751	7.136	7.288	-2.750	2.238
	(9.269)	(8.554)	(14.316)	(14.530)	(14.140)	(12.586)
Observations	728	728	258	258	470	470
Controls for parental background	-	✓	-	✓	-	✓

Notes: The presented numbers are the estimated effects of being born in July as opposed to being born in May using the 2006 PISA sample. The results in the top panel stem from the pooled sample of males and females, the middle panel from the male and the bottom panel from the female sample. All regressions include children born between May 1 and August 31 and control for birth months (dummy variables) and gender (top panel). The control variables on parental background include

background. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

6.2 Difference-in-Differences estimates of the effect of the parental leave reform on test scores

The results of the main DID estimates of the effect of the parental leave extension on PISA test scores are presented in Table 7 for the full sample (all mothers) and separately for mothers with higher and lower labour force participation rates. The estimation window includes children born two months before and after the reform (born between May and August). For each of these samples the first column shows the results without any control variables (columns 1, 3 and 5), while the second column contains results after controlling for background variables to account for potential changes in sample composition across years (and possibly across months). Even this more refined estimation approach seems to confirm the findings from the simple graphs shown earlier: the treatment effects of the reform (being born post June in the year 1990) are close to zero in the full sample (columns 1 and 2). Adding controls changes the coefficient slightly, but given the size of the standard errors these differences are not significant.[32] The average effect on reading test scores is a little bit larger and negative, but still not significantly different from zero.

In contrast to these neutral findings are the results from the regressions based on the subsample of high LFP mothers. The average effect on all three test subjects is significantly positive and of remarkable size; according to these results the extended parental leave period raised test scores by about 22 percent of a standard deviation (a little bit less in mathematics).[33] It is likely that these differences in results between the full sample and the sample of high LFP mothers can be related to the different levels of eligibility of the reform in conjunction with possibly higher quality care of better educated mothers (relative to informal care arrangements). Furthermore, the results for the subgroup of high LFP mothers are more likely to show the pure *time and care* effect of the parental leave extension, since the reform did not affect overall fertility or labour market success of mothers with higher earnings (which is proxied here by higher levels of education and higher labour force participation rates).

---

[32] Standard errors were clustered by school programme, school location and gender to account for the fact that test scores of students in the same programme, location and gender are likely to be correlated and not independent of each other.

[33] As a further refinement, a set of dummy variables for the different school programs (which are highly correlated with different average test scores) was added to the regressions. As expected, this inclusion helped to increase the precision of the estimates as the standard errors became smaller. The treatment effects became larger and more significant. However, due to the endogeneity of these variables these results are only reported in the appendix, Table A 2, column set (A).

In comparison, the effects for the low LFP group are generally negative and even statistically significant for the reading test scores. Although the fraction of affected mothers in this group is smaller, it seems that these children have not benefited from the reform. This could be potentially attributed to the increased levels of fertility and less available resources (time and market goods) per child (i.e. a possible quantity-quality trade-off), to shorten time intervals between births or maybe to a lower quality of maternal time in the sense of ability to foster cognitive child development.

To shed more light on potential gender differences regarding the effects of maternal employment on cognitive development of children found in other studies, the same analysis is repeated for boys and girls separately (Table 8 presents only the coefficients of the treatment effects of the reform).

Even though sample sizes become rather small in this subgroup analysis, the estimates seem to indicate that boys react more strongly to the reform than girls, especially when looking at the group of mothers with higher LFP rates. While both post-reform girls and boys have higher test scores on average, the coefficient for the boys is almost three times as large as that for girls. Furthermore, the effect remains only significantly different from zero for the male subgroup as the coefficients are too imprecisely estimated in the female subgroup (the effect on reading and science test scores for males corresponds to about 0.3 and 0.4 standard deviations). These results would be in line with potential differences in needs and in development between girls and boys at very young ages causing boys to benefit comparatively more from maternal care between the age of one and two. As before, the results for the pooled sample are rather small and insignificant and the results for the group of mothers with lower LFP rates have a negative sign.

Table 7: Difference-in-Differences estimation results (boys and girls)

	(1) Full sample	(2) Full sample	(3) High LFP mother	(4) High LFP mother	(5) Low LFP mother	(6) Low LFP mother
Mathematics						
Treatment effect	-0.406	2.000	17.158*	16.067*	-7.106	-5.021
	(6.855)	(6.698)	(9.685)	(9.177)	(9.554)	(8.829)
Post June	0.182	-2.541	-21.675*	-19.536*	8.308	5.954
	(7.008)	(6.768)	(10.889)	(9.904)	(8.577)	(7.985)
Born May	-2.335	-4.190	-11.698	-14.114	1.044	0.195
	(5.856)	(5.599)	(8.948)	(8.600)	(5.996)	(5.712)
Born July	-3.804	-2.946	8.361	5.116	-9.964	-8.397
	(5.193)	(4.754)	(8.805)	(8.625)	(6.691)	(5.151)
Year 2006	4.993	3.943	-6.723	-6.249	8.143	8.345
	(6.436)	(5.844)	(10.352)	(8.933)	(6.810)	(6.386)
Constant	503.812***	418.476***	528.176***	457.109***	494.827***	407.570***
	(12.363)	(18.962)	(14.228)	(23.328)	(12.398)	(20.766)
Reading						
Treatment effect	-7.362	-4.138	22.710*	21.220*	-19.811*	-17.144*
	(8.589)	(8.093)	(12.390)	(11.672)	(10.442)	(9.450)
Post June	4.555	1.302	-24.609*	-22.431**	15.431*	13.508*
	(7.596)	(7.279)	(12.399)	(11.074)	(7.930)	(7.473)
Born May	6.910	4.674	0.601	-2.670	8.447	8.182
	(4.948)	(4.701)	(10.120)	(8.886)	(5.723)	(5.763)
Born July	-1.769	-0.869	16.953	13.304	-11.251	-9.860
	(5.743)	(5.295)	(10.178)	(10.022)	(7.927)	(6.011)
Year 2006	6.476	4.842	-13.517	-12.230	12.997*	12.764*
	(6.958)	(6.423)	(11.796)	(11.183)	(7.518)	(7.140)
Constant	515.920***	408.674***	545.670***	447.464***	505.305***	399.338***
	(13.032)	(22.197)	(15.666)	(27.492)	(13.621)	(24.844)
Science						
Treatment effect	-1.009	2.103	23.713**	22.951**	-11.102	-8.454
	(7.623)	(7.411)	(10.329)	(10.020)	(9.939)	(9.041)
Post June	0.686	-2.749	-22.041**	-20.678**	8.964	6.426
	(7.004)	(6.814)	(10.603)	(9.660)	(8.579)	(8.011)
Born May	3.855	1.607	0.473	-2.685	4.232	3.384
	(5.944)	(5.677)	(7.782)	(6.878)	(6.900)	(6.757)
Born July	-1.083	-0.068	13.061	9.648	-8.223	-6.618
	(5.294)	(4.953)	(8.716)	(8.708)	(7.492)	(5.770)
Year 2006	22.148***	20.497***	5.236	5.398	27.570***	27.240***
	(6.812)	(6.154)	(10.186)	(8.797)	(7.330)	(6.902)
Constant	495.036***	388.403***	520.461***	419.676***	485.895***	380.347***
	(12.660)	(18.083)	(14.611)	(24.012)	(12.929)	(18.672)
Observations	2,840	2,840	943	943	1,897	1,897
Controls						
Parental background	-	✓	-	✓	-	✓

Notes: All regressions control for gender. The control variables on parental background include dummy variables for

parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table 8: Difference-in-Differences estimation results by gender.

	(1) Full sample	(2) Full sample	(3) High LFP mother	(4) High LFP mother	(5) Low LFP mother	(6) Low LFP mother
			MALES			
Mathematics	0.092	0.264	17.935	15.832	-8.420	-9.027
	(7.906)	(8.087)	(12.438)	(12.282)	(13.177)	(11.769)
Reading	-7.309	-6.808	34.009**	33.118**	-27.664*	-26.634**
	(9.794)	(9.750)	(15.154)	(14.985)	(14.516)	(12.868)
Science	-2.063	-1.232	41.033***	40.396***	-23.698	-23.251*
	(9.177)	(9.208)	(10.875)	(11.447)	(15.409)	(13.379)
Observations	1,426	1,426	482	482	944	944
			FEMALES			
Mathematics	-0.493	4.104	18.164	16.001	-5.617	-2.023
	(11.066)	(10.743)	(15.450)	(15.184)	(14.318)	(13.281)
Reading	-7.416	-1.890	13.368	13.905	-12.800	-8.914
	(13.684)	(12.585)	(19.744)	(19.082)	(15.466)	(13.978)
Science	0.561	5.762	7.778	6.330	1.670	5.818
	(12.139)	(11.878)	(16.723)	(15.817)	(14.109)	(13.126)
Observations	1,414	1,414	461	461	953	953
Controls						
Parental background	-	✓	-	✓	-	✓

Notes: The upper panel includes only male, the lower panel only female students. All regressions include dummy variable controls for survey year, birth months and for all children born post June. The control

attainment, school location, and migration background. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Given the trade-off between a larger sample size which could help to increase precision of the estimates on the one hand and the need to keep the estimation window as narrow as possible to reduce the influence of possibly confounding factors (the school-entry effect due to the proximity to the school entry cut-off date, age-at-test effect, seasonality), the regressions in Table 7 and Table 8 are repeated for wider samples which are gradually reduced to the more narrow window of four months (two pre- and two post-reform months) presented above (the widest sample includes children born between March and October).[34]

[34] Results using the narrowest window possible i.e. children born in June or July are reported in Table A 2, column set (B) in the appendix. These estimates are even larger and have a higher level of

The results for the symmetric extension of the sample (four/three/two months before and after the reform) are shown in Table 9 and are reported separately by maternal labour force participation subgroups (columns (3) and (6) correspond to the results from the two previous tables, Table 7 and Table 8). Looking at the first three columns for mothers with higher education it becomes evident that increasing the sample size symmetrically to the left and to the right of the cut-off leads to smaller coefficients; in other words, a narrower window around the cut-off date leads to a comparison of more similar the students resulting in larger effects. As a result and although the coefficients remain positive, the estimates in the top panel for the pooled male and female sample based on the wider window become insignificant. Nevertheless, the estimated effects for boys remain highly significant and positive, while those for girls get much smaller and even switch signs. Similarly, the results for the low LFP subgroup become less negative (or more positive) the more months are included in the analysis. Nevertheless, most of the results remain negative albeit often insignificant; the strongest effects are still found for boys.[35]

Table 10 summarizes further checks on whether these results are influenced by children born after the school-entry cut-off age (after September 1) and who are thus in lower grade levels than the other students. This time, additional months were added only to the left of the reform date, i.e. before July, while holding the number of months to the right constant (July and August). However, although the estimated coefficients are slightly larger (for the high LFP subgroup), the general picture remains the same.

---

significance than the estimates based on the four months sample, which could be driven by the omission of August children who are more likely to be in lower grade levels (and thus have lower test scores on average).

[35] Since the share of child toSenn(h)6(e0BT61.53 1 122.42 185.83 Tm[(-)] TJETBT1218.21 122.42 185.2-13(n)662-5(e

Table 9: DID estimates based on symmetrically extended estimation samples (up to four pre- and post-reform birth months)

	High LFP mothers			Low LFP mothers		
	(1) Mar-Oct	(2) Apr-Sep	(3) May-Aug	(4) Mar-Oct	(5) Apr-Sep	(6) May-Aug
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	5.678	9.944	16.360*	2.248	0.526	-5.634
	(7.713)	(7.315)	(9.274)	(6.483)	(7.313)	(9.140)
Reading	9.222	13.141	20.604*	-8.113	-9.716	-16.309*
	(9.455)	(9.811)	(11.724)	(6.831)	(7.656)	(9.245)
Science	11.661	15.918*	23.068**	-3.119	-5.116	-8.805
	(8.650)	(8.938)	(10.076)	(7.087)	(7.706)	(9.140)
Observations	1,887	1,425	943	3,772	2,840	1,897
	MALES			**MALES**		
Mathematics	13.328	13.487	15.832	-2.974	-3.577	-9.027
	(11.192)	(9.667)	(12.282)	(8.868)	(10.290)	(11.769)
Reading	27.162**	28.860**	33.118**	-16.769	-20.926*	-26.634**
	(11.567)	(12.027)	(14.985)	(10.094)	(12.214)	(12.868)
Science	32.614***	34.360***	40.396***	-16.917*	-18.967	-23.251*
	(10.420)	(9.864)	(11.447)	(9.709)	(11.123)	(13.379)
Observations	953	716	482	1,866	1,397	944
	FEMALES			**FEMALES**		
Mathematics	-0.186	5.977	16.001	7.340	4.872	-2.023
	(10.628)	(11.017)	(15.184)	(8.960)	(10.040)	(13.281)
Reading	-9.740	-1.316	13.905	3.965	2.321	-8.914
	(14.472)	(14.614)	(19.082)	(9.056)	(9.913)	(13.978)
Science	-8.190	-2.603	6.330	11.648	9.335	5.818
	(11.611)	(12.598)	(15.817)	(10.141)	(11.132)	(13.126)
Observations	934	709	461	1,906	1,443	953

Notes: The reported estimated treatment effects stem from separate estimations of different specifications. All regressions include dummy variables for month of birth, a year dummy for 2006, a dummy variable for all children born after June. Estimations from columns 1, 2, 4, and 5 include a further dummy variable for all children born between September and December to account for the school entry cut-off date in Austria and an interaction effect of this dummy variable with the year 2006 variable to account for potential general trends in school entry or repetition norms. The control variables on parental background

of the family. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table 10: DID estimates based on sample adding more pre-reform birth months while holding the number of post-reform birth months constant

	High LFP mothers			Low LFP mothers		
	(1) Mar-Aug	(2) Apr-Aug	(3) May-Aug	(4) Mar-Aug	(5) Apr-Aug	(6) May-Aug
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	6.037	10.427	16.360*	2.381	0.767	-5.634
	(7.720)	(7.383)	(9.274)	(6.534)	(7.356)	(9.140)
Reading	9.609	13.807	20.604*	-8.117	-9.529	-16.309*
	(9.380)	(9.694)	(11.724)	(6.964)	(7.756)	(9.245)
Science	12.066	16.600*	23.068**	-3.046	-4.961	-8.805
	(8.565)	(8.852)	(10.076)	(7.184)	(7.759)	(9.140)
Observations	1,407	1,178	943	2,830	2,367	1,897
	MALES			**MALES**		
Mathematics	13.612	13.594	15.832	-2.555	-3.134	-9.027
	(11.056)	(9.807)	(12.282)	(8.798)	(10.209)	(11.769)
Reading	27.569**	28.932**	33.118**	-16.415	-20.235	-26.634**
	(11.609)	(12.154)	(14.985)	(10.110)	(12.242)	(12.868)
Science	33.445***	35.290***	40.396***	-16.873*	-18.723	-23.251*
	(10.514)	(10.098)	(11.447)	(9.680)	(11.060)	(13.379)
Observations	718	594	482	1,424	1,174	944
	FEMALES			**FEMALES**		
Mathematics	0.947	6.821	16.001	7.340	4.731	-2.023
	(10.881)	(11.234)	(15.184)	(9.294)	(10.293)	(13.281)
Reading	-8.025	0.064	13.905	4.100	2.223	-8.914
	(14.513)	(14.782)	(19.082)	(9.328)	(10.085)	(13.978)
Science	-6.790	-1.662	6.330	12.025	9.341	5.818
	(11.657)	(12.657)	(15.817)	(10.495)	(11.267)	(13.126)
Observations	689	584	461	1,406	1,193	953

Notes: The reported estimated treatment effects stem from separate estimations of different specifications. All regressions include dummy variables for month of birth, a year dummy for 2006, a dummy variable for all children born after June. The control variables on parental background include dummy variables al attainment, school location, and migration background of the family. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

6.3 Difference-in-Difference-in-Differences estimations (DDD)

The DID estimation strategy assumes that the age-related differences across birth months (season-of-birth, age-at-test, school-entry effect) are constant over time or at least over the period of three years that lies between the two PISA waves (the treatment cohort tested in 2006 and the control cohort tested in 2003). As a further check and control for possibly confounding changes in relative birth month outcomes, an additional control group will be included in the analysis that participated in the PISA test in both relevant years, but was not affected by the parental leave reform in 2006. ouring country Germany has a very similar schooling and tracking system (Schneeweis and Zweimüller 2009) and both countries are also very close in terms of their child care institutions as well as cultural values and attitudes towards the role of families and mothers (Neyer 2003). Furthermore, the PISA test language in both countries is German and thus Germany seems to be a suitable candidate to act as additional control group in the analysis.

By including a further control group to the analysis, the effect of the parental leave reform on child outcomes will be estimated using the following triple difference estimation specification, which adds another country variable, as well as interaction effects between this country variable and Post June births and the year dummy variable 2006 and the triple interaction term of the Post June, Year 2006 and Austria dummy variables.

$$\boldsymbol{y_i = \alpha + \beta_1 Post\,June + \beta_2\, y2006 + \beta_3 Post\,June * y2006 + \beta_4 Austria + \beta_5 Austria * y2006 + \beta_6 Austria * Post\,June + \beta_7 Post\,June * y2006 * Austria + \theta_m birth\,month + \mu X + \varepsilon_i} \tag{4.5}$$

The OLS estimate of the treatment effect $\boldsymbol{\widehat{\beta}_7}$ now becomes (AUT stands for Austria (treatment country, in which the reform takes place); GER stands for Germany (control country)):

$$\boldsymbol{\widehat{\beta}_7} = [(\boldsymbol{\bar{y}_{2006,AUT,post}} - \boldsymbol{\bar{y}_{2006,AUT,pre}}) - (\boldsymbol{\bar{y}_{2006,GER,post}} - \boldsymbol{\bar{y}_{2006,GER,pre}})] - [(\bar{y}_{2003,AUT,post} - \bar{y}_{2003,AUT,pre}) - (\bar{y}_{2003,GER,post} - \bar{y}_{2003,GER,pre})]. \tag{4.6}$$

Table 11 displays the estimated treatment effect of the parental leave reform in Austria based on the DDD regressions. Generally the DDD results correspond to the DID estimates. Children of mothers with higher labour force participation, seem to have

benefited from the parental leave extension the estimated coefficients are positive for all children, but only significant and larger for boys (the estimated coefficients for males are slightly larger than the DID estimates corresponding to approximately between 0.4 and 0.7 standard deviations). Although the average estimated effects of the pooled sample of males and females are slightly larger than the DID estimates, they are slightly less significant for the May to August sample. Hence, although the additional observations from Germany help to almost double the sample size, they do not help to increase the precision of the estimation. In contrast, a comparison of the standard errors across the wider and narrower samples (March-Octob7[(lu.t[(O)-7(uss)-18Ms)-10nay

i o n d i n g - 9 (   ) 6 2 9 (f o r) 6 (   ) 5 9 7   b o y   (O) - 7 (e   ) 5 1 7 (w t h e   ) 6 1 [l (o h) - 9 (e) 4 (   ) -

ofr(inlucde)4(dinm)-11 (theMa)6n(a)4( )-16(y)20(isl)-3

p a r g e . r m a n   s l i g n   f i c a n s) ;   r e g   t h e   f r e c s (l u . 2 [(f r) - 6 r   r) 1 -

t   o w e   a e

t e e w e f f e c (s   )]   T J   E T   B T   1   0   0   1   6 5 . 9 8 2   7 . 7 3 0 4   T m   [(f o r) 6 ((t h e   ) - 7 (g) 1 0

Table 11: DDD estimations including German students as further control group

	High LFP mothers			Low LFP mothers		
	(1) Mar-Oct	(2) Apr-Sep	(3) May-Aug	(4) Mar-Oct	(5) Apr-Sep	(6) May-Aug
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	19.210*	22.205*	22.765	-6.174	-9.851	-17.131
	(9.922)	(11.276)	(14.804)	(10.027)	(11.674)	(14.101)
Reading	18.640	20.605	25.079	-23.189*	-26.549*	-30.991*
	(13.503)	(15.201)	(17.315)	(12.286)	(13.369)	(15.499)
Science	29.104**	33.547**	37.808**	-14.867	-16.967	-20.919
	(11.817)	(13.318)	(15.267)	(11.326)	(13.141)	(14.809)
Observations	4,228	3,212	2,158	6,519	4,968	3,325
	MALES			**MALES**		
Mathematics	28.476**	32.697**	38.109**	-33.133**	-34.056*	-37.219**
	(12.966)	(13.427)	(18.223)	(15.778)	(17.211)	(16.667)
Reading	32.198*	36.062*	40.010*	-55.182***	-56.815**	-57.088**
	(16.737)	(18.282)	(22.573)	(19.308)	(21.203)	(22.137)
Science	48.607***	56.602***	66.932***	-51.510***	-53.240***	-53.071***
	(10.574)	(12.055)	(15.812)	(16.249)	(18.578)	(18.600)
Observations	2,157	1,634	1,113	3,209	2,438	1,636
	FEMALES			**FEMALES**		
Mathematics	11.050	11.602	9.412	16.207	10.129	2.579
	(16.350)	(18.952)	(23.564)	(11.318)	(15.010)	(21.739)
Reading	0.896	3.155	13.632	7.913	1.553	-8.546
	(18.381)	(21.168)	(21.205)	(12.644)	(16.166)	(23.939)
Science	9.298	9.523	10.392	16.910	14.040	9.852
	(18.943)	(21.459)	(23.252)	(13.299)	(17.327)	(22.128)
Observations	2,071	1,578	1,045	3,310	2,530	1,689

Notes: The reported estimated treatment effects are from separate estimations of different specifications. All regressions include dummy variables for month of birth, year and country fixed effects, a dummy variable for all children born after June, interaction effects between year and the post June dummy, year and country, country and post June . Estimations from columns 1, 2, 4, and 5 include a further dummy variable for all children born between September and December to account for the school entry cut-off date in Austria (and a year interaction). The control variables on parental background include dummy tion background of the family. Robust standard errors in parentheses (clustered by school track (more/less academic)[36], school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

[36] Since Austrian and German school types are not fully comparable, an alternative, comparable school type measure was constructed which differentiates between schools that do or do not provide access to university or college after graduation (academic track versus non-academic track).

7 Further robustness checks

This section presents the findings from several alternative estimations, placebo tests and sensitivity checks in order to test the robustness of the previous results. To reduce the extensiveness of the tables, only results from the regressions by gender subgroup will be presented: the question is whether the previously found strong effects for boys prove robust.

7.1 DID using German students as alternative control group

A first alternative estimation strategy is to use only data from the PISA test 2006 and to implement a DID estimator using German students as a control group (instead of Austrian students from the pre-reform year 2003). This estimation strategy replaces the assumption of a common birth month trend across years by the assumption of a common birth month trend across regions. Again, the results for the children of the high LFP group of mothers correspond to the previous findings (see Table 12; although the coefficients are slightly larger than in the original DID regressions and the effects on test scores in mathematics for boys become significant). For the group of mothers with lower LFP rates the effects remain negative, but insignificant for boys. However, the coefficients for the female subgroup of less educated mothers increase in size and in the specifications using the extended sample (i.e. children born after the Austrian school-entry cut-off date) the coefficients become even significantly positive. Nevertheless, these positive effects are much smaller and less significant when using only the standard four months window (May-August). In contrast to Austria, the school-entry cut-off date in Germany is determined separately in each of the 16 federal states and thus varies across the country.[37] Still, in many federal states the cut-off date is June 30 (Keil 2005), which would coincide with the reform cut-off date. This could potentially lead to a violation of the common trend assumption if this causes a drop in average test scores around this threshold in the German sample. While the DDD estimates accounted for this potential drop by differencing across years, this might be a problem in the DID regression set up (by artificially raising the treatment effect).

[37] Optimally, one could have limited the German sample to those regions, which are culturally closest to Austria (i.e. Bavaria, which shares a common border with Austria). Unfortunately, the international PISA data files do not contain regional identifiers.

Table 12: Robustness check. DID estimations using only PISA 2006 data and Germany as a control group.

	High LFP mothers			Low LFP mothers		
	(1) Mar-Oct	(2) Apr-Sep	(3) May-Aug	(4) Mar-Oct	(5) Apr-Sep	(6) May-Aug
		MALES			**MALES**	
Mathematics	22.887*	31.897**	34.045**	6.172	1.653	-10.321
	(11.024)	(11.212)	(14.802)	(13.336)	(15.374)	(15.686)
Reading	31.732*	40.370**	44.189*	-2.666	-7.632	-15.874
	(16.611)	(17.236)	(22.632)	(16.095)	(17.192)	(17.525)
Science	34.776**	44.846***	51.963***	-4.358	-10.952	-21.228
	(13.091)	(12.187)	(16.236)	(11.271)	(12.525)	(12.549)
Observations	1,202	911	618	1,621	1,226	826
		FEMALES			**FEMALES**	
Mathematics	15.409	15.938	21.866	22.334***	17.954**	11.546
	(16.665)	(17.766)	(18.373)	(5.781)	(8.156)	(11.338)
Reading	5.420	4.990	12.259	14.473**	9.908	8.082
	(14.225)	(15.269)	(15.076)	(6.885)	(9.419)	(12.153)
Science	8.694	9.160	13.475	20.077**	18.477*	19.073*
	(14.292)	(15.279)	(15.227)	(7.969)	(9.619)	(10.982)
Observations	1,151	888	596	1,675	1,261	843

Notes: The reported estimated treatment effects are from separate estimations of different specifications. All regressions include dummy variables for month of birth, a dummy variable for all children born after June. Estimations from columns (1) (4) include a further dummy variable for all children born between September and December to account for the school entry cut-off date in Austria. The control variables on

migration background of the family. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

7.2 Placebo tests

In this subsection the results of three placebo tests are presented: First, a DID regression using exclusively the German subsample and treating July 2006 as pseudo reform month (column set (A) in Table 13), second, a DID regression using only data from the non-reform year 2003 for Austria and Germany, treating July 2003 in Austria as pseudo reform month (column set (B) in Table 13), third, using the original DID estimation strategy based on Austria and the years 2006 and 2003, but using alternative months as pseudo cut-off dates of the reform, namely May 2006 and June 2006 (Table 14).

Table 13: Placebo tests using the German subsample

	(A) DID using Germany only (2006 and 2003)		(B) DID using 2003 only; (Austria and Germany)	
	High LFP mothers	**Low LFP mothers**	**High LFP mothers**	**Low LFP mothers**
	(1) May-Aug	(2) May-Aug	(3) May-Aug	(4) May-Aug
	MALES		**MALES**	
Mathematics	-18.912	28.261**	-5.108	27.613**
	(13.572)	(9.814)	(19.843)	(9.970)
Reading	-4.714	28.811	4.011	40.045***
	(13.985)	(16.458)	(18.768)	(11.639)
Science	-24.595*	28.987**	-15.874	31.872**
	(13.076)	(12.100)	(18.373)	(11.878)
Observations	631	692	495	810
	FEMALES		**FEMALES**	
Mathematics	4.981	-3.058	15.117	7.042
	(16.789)	(10.460)	(21.769)	(14.299)
Reading	-1.027	-0.012	-0.965	13.276
	(11.016)	(8.516)	(21.035)	(16.328)
Science	-6.044	-3.224	4.422	8.101
	(13.003)	(11.323)	(21.984)	(14.966)
Observations	584	736	495	846

Notes: Each cell reports the treatment effect estimated using different specifications. All regressions include dummy variables for month of birth and a dummy variable for all children born after June as well as controls for
The regressions in the left two columns include furthermore a year dummy as well as year-post June interaction effect. The two columns on the right contain a country variable for Austria and an Austria-post June interaction effect. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

The results of the first test based on PISA data for Germany only are mostly insignificant for children of mothers with higher education as they should be given the pseudo reform assumption of the test (Table 13). Only the coefficient for the science test scores for the male sample is negative and significant at the 10 percent significance level. What is more worrying are the significantly positive effects of the pseudo treatment on maths and science test scores for sons of mothers with lower LFP rates (the coefficients for the girls are very small and insignificant). These results indicate that for

this particular group there is a positive trend in Germany over time and this could possibly lead to an artificial inflation of the negative coefficients in the DDD regression framework. The second test, using Austrian and German PISA data from the non-reform year 2003 again shows no significant effects for the group of children with high LFP mothers. This result gives more credibility to the positive effects for boys found in the DID regressions using Austrian and German data from 2006 (see Table 12). Furthermore, these results do not seem to support the hypothesized problem of the German school-entry cut-off dates. On the other hand, the results for boys of the low LFP group are disturbing: there seems to be an increase in test scores for this particular group in this regression comparing Austria with Germany which again could bring about the stronger negative results in the DDD framework. Hence, as long as the high LFP group is concerned these two placebo tests using German data tend to support the validity of the previous results and the use of the German data as additional control group. Regarding the results for children from the lower LFP group it seems that there might be confounding trends in Germany, which casts doubts on the reliability of the DDD estimates for this subgroup.

The third placebo test using alternative pseudo reform dates is presented in Table 14. The three columns in column set (A) come from DID regressions using May 1, 1990 as pseudo reform date and compares the test scores of children born two months prior and after the pseudo reform cut-off date (born March-April versus May-June). None of the estimated effects is significantly different for the male subsample (irrespective of maternal labour force participation rates). As regards the female subsample, all coefficients except for one are not significantly different from zero (the exception is a significantly negative effect for the reading score in the group of mothers with high LFP rates). Thus, overall, but especially regarding the male subsample, it seems that the placebo test supports the idea that the DID effect found in the original regressions are not driven by statistical outliers or seasonal/yearly patterns or trends.

Table 14: Placebo tests using other months as pseudo reform dates

	(A) Pseudo cut-off is *May 1, 1990* DID 2006+2003			(B) Pseudo cut-off is *June 1, 1990* DID 2006+2003		
	Full sample	**High LFP mothers**	**Low LFP mothers**	**Full sample**	**High LFP mothers**	**Low LFP mothers**
	(1)	(2)	(3)	(4)	(5)	(6)
	Mar-Jun	Mar-Jun	Mar-Jun	May-Jun	May-Jun	May-Jun
		MALES			**MALES**	
Mathematics	4.007	-5.903	10.418	-39.341***	-69.630***	-29.584*
	(10.081)	(14.688)	(12.885)	(12.065)	(19.648)	(15.088)
Reading	7.702	-14.833	19.133	-33.873**	-69.286**	-23.999
	(11.491)	(17.469)	(14.598)	(14.916)	(26.523)	(18.220)
Science	1.185	-15.328	10.248	-38.284**	-58.028**	-34.799*
	(11.928)	(13.059)	(15.292)	(17.216)	(23.587)	(19.251)
Observations	1,432	462	970	716	226	490
		FEMALES			**FEMALES**	
Mathematics	1.437	-26.332	18.209	-16.537	-33.782	-12.179
	(11.353)	(15.491)	(16.229)	(15.080)	(25.184)	(15.695)
Reading	2.017	-39.974**	25.166	-13.607	-23.318	-13.127
	(13.365)	(15.263)	(19.199)	(13.753)	(25.679)	(13.450)
Science	-1.225	-23.605	12.143	0.846	-9.598	3.672
	(9.896)	(13.953)	(12.958)	(14.874)	(31.326)	(14.183)
Observations	1,361	450	911	680	222	458

Notes: Each cell reports the treatment effect estimated using different specifications. All regressions

location, and migration background of the family. The three left columns include furthermore control variables for birth months, a dummy variable for all children born after the pseudo cut-off date May 1, 2006 (May and June) as well as year-post April interaction effect. The three columns on the right contain a dummy variable for all children born after the pseudo cut-off date June 1, 2006 and an interaction of this dummy with the year dummy 2006. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

However, turning to the three columns on the right which treat June 2006 (the month prior to the actual reform) as pseudo reform date and including only one pre and post pseudo-reform months (May and June) the picture looks less promising. Although the placebo test would require all coefficients to be insignificant, the estimated pseudo-reform effects are significantly negative and quite large for the male subgroup (although the samples become extremely small) and negative but insignificant for the girls. This finding is puzzling and the question is, whether June 2003 or June 2006 are outliers in the data who drive the original effects (if there is a negative June effect, than the

estimated positive treatment effect of the reform on July 1 for sons of higher LFP mothers might be driven by this negative pre-reform effect). The significant June effects are worrying and require a further repetition of the original DID estimates *excluding* children born in June to test to what extent the June effect seems to be driving the result (see next section).

Generally, it could also be that June 2003 is an *outlier* and it would be helpful to test this possibility by repeating the DID analysis using another control year with PISA test scores from Austria instead of PISA 2003. Alternative robustness checks could have involved data from the Austrian PISA tests 2000 and 2009 (i.e., birth cohorts 1984 and 1993). However, several problems with data from these two years made them unsuitable for this purpose.[38] First of all, the sample of schools in included in the sample frame in the year 2000 was biased, since students enrolled in combined school and work vocational programmes were systematically underrepresented. Furthermore, substantial revisions of the student background questionnaire between 2000 and 2003 make it impossible to construct comparable and consistent categories of parental education. As regards the 2009 data,

> *"a dispute between teacher unions and the education minister in Austria led to the announcement of a boycott of PISA which was withdrawn after the first week of testing. The boycott required the OECD to remove identifiable cases from the dataset. Although the Austrian dataset met the PISA 2009 technical standards after the removal of these cases, the negative atmosphere in relation to education assessments affected the conditions under which the assessment was administered and could have adversely affected student motivation to respond to the PISA tasks"* (see footnote 35).

An even bigger problem that impedes the inclusion of the 2009 data into the analysis comes from the fact that the cohort of included children (born in 1993), were the first cohort to be affected by stricter school entry rules (entry into first grade) that came into effect at the beginning of the school year 1999/2000. An amendment to the school law caused an increase in compliance with the school-entry cut-off date regulations. As a consequence the rate of children entering school late became much smaller. It is possible that these stricter school entry regulations led to an increase of repetition rates for children, who would have otherwise started school one year later. These changes and consequences affected mainly children born shortly before the cut-off date of September 1, i.e. born in August or July. Thus, there is a different trend over

[38] See also the information on Anomalies in PISA data available on the OECD web-site, http://www.pisa.oecd.org/document/53/0,3746,en_32252351_32235731_38262901_1_1_1_1,00.html.

time regarding the relative composition of July and August children in terms of age at school entry and subsequent school experiences in comparison with children born in other months of the year.

7.3 DID estimates excluding June children

As explained above, it is possible that the estimated treatment effect found in the earlier regressions is driven by an unexpectedly negative trend in test scores of June children (while these could be driven either by June 2003 test scores being unnaturally high *or* June 2006 test scores being unnaturally low). This is why the previous DID estimations are repeated without including June children in the estimation sample. Certainly, given the rather small size of the estimation sample, dropping observations from an entire birth months affects the precision of the estimates. Furthermore, the post-reform results including the potentially problematic August and September children due to the school-entry effect receive relatively more weight. Table 15 shows the main results for the different sample specifications (symmetric extension to the left and to the right of the reform cut-off, as well as including only July and/or August as post-reform months) and also for the DID estimates using only 2006 and German students instead of Austrian students from 2003 as control group (the two columns on the right of the table).

Once June is excluded from the sample, the estimated coefficients for children from the group of mothers with higher maternal LFP rates become much smaller in size and mostly insignificant. In fact, the estimated treatment effects become negative for girls across all subjects and for boys with respect to mathematics test scores. Nevertheless, the estimated treatment effects on reading and science for the male subsample remain positive and are partly significantly different from zero on the ten percent level (column two and four). The estimated treatment effects for children whose mothers have a lower LFP rate become more negative. In contrary, the re-estimated DID estimates based on the 2006 data with Germany as control region remain significantly positive for boys of mothers with higher LFP rates (columns eleven and twelve).

Table 15: DID estimates excluding children born in June from the regressions

	DID estimates (Austria 2006/2003)										DID 2006 estimates (Austria/Germany)	
	High LFP mothers					**Low LFP mothers**					**High LFP mothers**	**Low LFP mothers**
	(1) Apr-Sep	(2) Apr-Aug	(3) May-Aug	(4) Apr-Jul	(5) May-Jul	(6) Apr-Sep	(7) Apr-Aug	(8) May-Aug	(9) Apr-Jul	(10) May-Jul	(11) May-Aug	(12) May-Aug
MALES												
Mathematics	-3.562	-3.053	-11.802	3.795	-5.300	-7.246	-6.937	-22.481	-6.216	-20.677	41.468*	-18.153
	(12.295)	(12.512)	(18.361)	(15.792)	(20.934)	(12.541)	(12.471)	(13.925)	(12.390)	(15.499)	(22.332)	(23.099)
Reading	11.644	12.055	6.050	30.448	25.624	-23.291	-22.913	-37.972**	-15.547	-29.394*	50.809*	-33.143
	(12.504)	(12.905)	(17.999)	(17.910)	(22.745)	(14.776)	(14.860)	(13.744)	(13.800)	(14.822)	(27.692)	(25.717)
Science	18.735	20.302*	17.775	29.408*	26.449	-24.883*	-24.807*	-39.587**	-27.274**	-40.945**	55.520**	-33.025*
	(11.360)	(11.718)	(16.363)	(14.842)	(19.116)	(13.158)	(13.097)	(15.462)	(12.149)	(16.484)	(21.224)	(18.487)
Observations	618	496	384	359	247	1,154	931	701	707	477	476	633
FEMALES												
Mathematics	-7.214	-6.956	-5.076	-11.251	-10.699	5.338	5.207	-7.976	3.350	-10.154	7.875	13.223
	(13.109)	(13.306)	(20.421)	(31.011)	(34.851)	(9.369)	(9.538)	(12.049)	(15.356)	(18.596)	(20.875)	(9.925)
Reading	-14.864	-13.951	-1.744	-16.931	-5.742	5.035	4.971	-15.054	0.264	-22.413	-6.757	-5.276
	(15.295)	(15.410)	(24.619)	(25.117)	(33.059)	(10.851)	(10.917)	(16.558)	(17.410)	(23.475)	(17.222)	(15.172)
Science	-10.327	-9.636	-1.957	-14.274	-7.914	12.026	12.053	7.861	12.846	6.944	-10.257	17.605
	(16.195)	(16.227)	(25.967)	(24.400)	(32.539)	(11.488)	(11.535)	(14.290)	(18.790)	(22.166)	(16.709)	(11.550)
Observations	604	479	356	354	231	1,201	951	711	691	451	461	655

Notes: Each cell reports the estimated treatment effects from separate DID regressions controlling for the standard set of background variables as in the original regressions in section 6. Children born in June are excluded from the estimation sample. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

To sum up, omitting June children from the estimation reduces the estimates. The previously positive treatment effects for boys from better educated mothers become insignificant in most specifications. Since these regressions gave relatively more weight to the post-reform birth months, the remaining question is to what extent the findings (including and excluding June) are affected by the proximity to the school-entry cut-off date and the corresponding higher fraction of students in lower grade levels in the post-reform months. This will be analysed in the following section.

7.4 DID estimates of other schooling outcomes and the role of retained students

To better understand the underlying mechanisms of the previous findings as well as the potentially confounding role of the school-entry cut-off age effect two further analyses are conducted. The first set of regressions repeats the DID estimations using alternative schooling outcomes, namely whether a student is in a lower than regular grade level given his birth month and whether a student is enrolled in the academic track. The research question motivating these regressions is whether the reform had an effect on these schooling outcomes as well. If so, the estimated treatment effect of the reform based on test scores might partly be driven by these mechanisms and the consequential grade composition of students for each birth month (more or less regular or retained students in the post-reform months July and August). Unfortunately, the data do not allow disentangling the reasons *why* students are in a lower than regular grade level. This could be either due to deferred school entry or to grade repetition.

The regression results for the effect of the extended maternal leave period on the two alternative schooling outcomes are reported in Table 16. Overall, the effects for both outcomes are small and in most of the cases not significantly different from zero. Nevertheless, there seem to be differences between the two groups of mothers: for the group of *mothers with lower LFP rates* (column 3) the results suggest that the extended leave period increased the likelihood that treated children are significantly more likely to be in a lower than regular grade level. This effect is significantly positive in the pooled male and female sample, but seems to be even more pronounced for girls than for boys. In line with these findings the coefficients of the reform on the propensity to

be enrolled in the academic track are all negative. However, none of these estimates is significant.

Table 16: Probability of being in lower than regular grade level (grade retention) or being enrolled in the academic track (linear probability models)

	DID 2006 and 2003 Two months window May August		
	(1) Full sample	(2) High LFP mothers	(3) Low LFP mothers
		MALES + FEMALES	
Retained	0.055	-0.022	0.089**
	(0.035)	(0.059)	(0.041)
Academic track	-0.042	0.004	-0.067
	(0.034)	(0.059)	(0.043)
Observations	2,840	943	1,897
		MALES	
Retained	0.019	-0.031	0.042
	(0.051)	(0.090)	(0.049)
Academic track	-0.044	-0.005	-0.064
	(0.036)	(0.062)	(0.047)
Observations	1,426	482	944
		FEMALES	
Retained	0.089**	0.002	0.137**
	(0.040)	(0.077)	(0.062)
Academic track	-0.047	0.019	-0.065
	(0.057)	(0.091)	(0.073)
Observations	1,414	461	953

Notes: Each cell reports the estimated treatment effects from separate DID regressions controlling for the standard set of background variables as in the original regressions in Section 6. The dependent variables are two dummy variables indicating whether a student is enrolled in a lower than regular grade level given his birth date and whether the student is enrolled in the academic track. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by

negative effect of the parental leave reform on the likelihood to be retained and a positive effect on being enrolled in an academic track school (none of the results for this group of mothers is significant though).

There also appear to be some differences between boys and girls which however seem to be driven by the subgroup of girls from mothers with lower LFP rates. For the boys, all of the estimated effects are statistically significant and close to zero. The same is true for daughters of mothers with higher LFP rates (no significant effects). However, in the full sample (all mothers) of girls the results seem to suggest a positive and statistically significant effect of the reform on the likelihood to be retained. This result seems to be driven by the group with low LFP mothers though.

To sum up, although the estimated coefficients are suggestive, they are all close to zero and mainly not significant. Most importantly, for the group of children from mothers with higher LFP rates there seems to be no effect of the reform on the two outcomes. In the full sample estimation for females as well as for the subgroup of mothers with lower LFP rates, there seems to be an effect on the likelihood to be in a lower than regular grade level. This could explain the respective negative findings in the test score regressions.

The next test repeats the DID estimations for the test score outcomes from Section 6.2, but limits the estimation sample to students who are not retained but currently enrolled in the regular grade level according to their birth month and the school-entry cut-off regulations. This will generally increase the average test scores in each birth month, since students in lower grades have lower test scores on average as they have less years of formal schooling and have covered less material in school. In addition, this will probably reduce the variance of test scores in each birth month (especially in the months close to the school-entry cut-off where there are more children on lower grade levels).

The first six columns in Table 17 are based on regressions *including* the potentially problematic month of June, while the six columns on the right side of the table are based on regressions *excluding* children born in June. Certainly, if the *outcome* variable of the reform, it is problematic to split the sample according to this variable.

However, the previous results have shown that there is no significant reform effect on the probability of grade retention for children of mothers with higher LFP rates and on boys; the exception are daughters of mothers with lower LFP and, hence, the results for this subgroup have to be interpreted cautiously.

The results in the first six columns from Table 17 can be compared to the results from Table 15 as these are estimated using the same estimation windows (samples with gradually narrowed birth month estimation windows). The general pattern of the estimated treatment effects corresponds well to the one using the full sample. However, the positive treatment effects for sons from mothers with higher LFP rates are slightly larger and slightly more significant, while the negative effect on sons of mothers with lower LFP rates become slightly weaker and less significant. The results for the female subsamples change only marginally, but remain insignificant. What is more interesting: repeating this set of estimations *excluding* children born in June does not affect these findings dramatically. The effect of the parental leave extension remains positive and large (and is significantly different from zero in many specifications) for sons of mothers with higher LFP rates. The estimated effects for the girls are not significantly different from zero through all specifications and for all subjects. On the other hand, the negative effect of the reform on children from mothers with lower LPF rates becomes slightly larger and partly significant; however, only in the May to August specification do the seem to be significantly affected by the reform.

Hence, although the parental leave reform has only negligible and insignificant effects on the likelihood of being retained for the subsample of children with mothers with higher LFP rates (as shown in Table 16), the restriction of the sample to regular students seems to reduce the sensitivity of the results with respect to the exclusion of June. However, if anything, boys with higher LFP mothers seem to be less likely to be retained as a result of the reform (even though the coefficient is not significant), meaning that restricting the sample to regular students should actually work in the other direction (making the reform effect less positive). Overall, even though it could be critical to restrict the sample to regular students, it helps to get an impression of how grade retention of students close to the school-entry cut-off date affect the DID results (namely through a possible downward bias).

Table 17: DID estimates using only students in regular grades (including and excluding June births)

	Regressions based on regular students, *including June*						Regressions based on regular students, *excluding June*					
	High LFP mothers			Low LFP mothers			High LFP mothers			Low LFP mothers		
	(1) Mar-Oct	(2) Apr-Sep	(3) May-Aug	(4) Mar-Oct	(5) Apr-Sep	(6) May-Aug	(7) Mar-Oct	(8) Apr-Sep	(9) May-Aug	(10) Mar-Oct	(11) Apr-Sep	(12) May-Aug
MALES												
Mathematics	22.924	25.806*	34.498*	3.848	3.164	-1.665	12.712	10.751	14.553	1.114	-1.890	-18.257
	(14.680)	(12.610)	(17.440)	(8.758)	(10.706)	(13.062)	(16.200)	(14.773)	(22.068)	(8.979)	(12.115)	(11.744)
Reading	37.448***	40.073***	46.034**	-13.098	-17.583	-22.845	27.349**	26.243**	26.485	-14.438	-21.957	-37.660**
	(8.680)	(8.678)	(16.907)	(12.289)	(14.401)	(15.197)	(10.483)	(10.880)	(20.688)	(12.810)	(16.652)	(13.297)
Science	49.459***	54.125***	66.762***	-9.099	-11.633	-16.041	40.195***	41.450***	51.701**	-12.028	-17.283	-33.328**
	(11.867)	(10.108)	(15.647)	(9.836)	(12.252)	(15.431)	(13.594)	(12.075)	(21.050)	(9.551)	(12.645)	(12.222)
Observations	704	504	305	1,496	1,076	707	637	437	238	1,308	888	519
FEMALES												
Mathematics	-2.633	4.970	12.703	7.594	2.813	3.047	-12.112	-4.003	-1.171	6.050	-1.977	-7.861
	(17.865)	(17.758)	(21.009)	(8.225)	(9.050)	(12.725)	(20.275)	(20.296)	(25.229)	(8.215)	(8.525)	(11.914)
Reading	-10.685	-1.671	13.178	2.323	-0.630	-2.665	-20.881	-12.360	3.265	1.105	-3.845	-12.066
	(20.900)	(20.847)	(25.449)	(9.911)	(10.597)	(15.513)	(21.133)	(20.729)	(27.753)	(9.930)	(10.229)	(17.009)
Science	-4.932	1.055	9.103	10.972	5.710	9.090	-10.736	-4.429	4.575	11.294	3.445	7.326
	(19.124)	(20.332)	(22.878)	(11.378)	(12.151)	(13.979)	(21.303)	(23.818)	(31.441)	(11.835)	(13.492)	(17.153)
Observations	746	555	351	1,576	1,162	740	661	470	266	1,379	965	543

Notes: Each cell reports the estimated treatment effects from separate DID regressions controlling for the standard set of background variables as in the original regressions in Section 6. The estimation sample includes only children in regular grade levels (according to their birth months). Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

8 Conclusions

The objective of the empirical analysis presented in this paper was to investigate whether a substantial extension of the paid and job-protected Austrian parental leave mandate had any medium-term effects on the cognitive development of children. What makes the Austrian parental leave reform particularly suitable for a causal Difference-in-Difference (DID) analysis is that it was implemented with a strict and unanticipated cut-off date: only those mothers who gave birth to their child on or after July 1, 1990 became eligible for the more generous 24 months parental leave duration. As a consequence of this unexpected cut-off date the allocation of mothers and their children into treatment (24 months parental leave entitlements) and control

the problem of the otherwise endogenous return-to-work decision. Another advantage of this particular Austrian reform is that it did not seem to have any effects on medium or long-term labour market outcomes of mothers and only a small positive effect on fertility (as shown by previous studies).

To assess the effect of the parental leave extension on child outcomes the empirical analysis made use of mathematics, reading and science test scores from the standardized PISA test (using the cohort born in the year of the reform 1990 as well as a control birth cohort which was not subject to a reform, 1987; the corresponding PISA tests were conducted in the years 2006 and 2003). Since the data do only contain information on maternal and paternal educational attainment, but neither on parental leave eligibility nor on actual leave taking, the estimated effects resemble intention-to-treat estimates (net reduced form effects).

The results of the DID analyses and several robustness checks reveal that there seem to be heterogeneous effects of the parental leave reform on PISA test scores across estimation samples and subgroups. When using the *full sample*, the estimates suggest that the 12 months parental leave extension did not have any statistically significant causal medium-run effects on cognitive skills. This finding for Austria is in line with the results of most of the studies using changes in parental leave mandates to identify the causal effect of early maternal employment on child outcomes (in particular, Canada

(Baker and Milligan 2010a, b), Denmark (Würtz Rasmussen 2010), Germany (Dustman and Schönberg 2010) and Sweden (Liu and Nordstrom Skans 2010)).

However, when splitting the sample into two groups of mothers with higher and lower labour force participation rates respectively (i.e., higher and lower educational attainment), interesting findings emerge: for the children of mothers with higher labour force participation rates and who are thus more likely to have been eligible and affected by the reform, there appears to be a significantly positive effect of the parental leave extension on the PISA test scores at age 15. This positive result seems to be especially driven by the large and significant effect on boys. Such a positive reform effect among the group of more eligible mothers would generally correspond to the findings by Carneiro, Løken and Salvanes (2010) for Norway: while their analysis using the full sample produces insignificant results, the estimated effects using the subsample of eligible mothers (which they can identify in the data) reveals significantly positive

Liu and Nordstrom Skans (2010) also find positive effects for their assessed parental leave extension in Sweden when restricting the sample to children of mothers with higher education. The other studies do either not test or simply do not find any differences according to maternal parental leave eligibility or educational status. The same is true for gender differences: while there exists some evidence of heterogeneous *health* outcomes of girls and boys with respect to early maternal employment, the previous evaluation studies of parental leave on *cognitive* child outcomes have not tested or not found corresponding gender effects. Against this background, the presumably stronger effect of the Austrian parental leave extension on boys points to an interesting aspect which should be explored in more detail in future research.

Although these empirical results are robust to various sensitivity checks including Difference-in-Difference-in-Differences estimations using Germany as an additional control region which was not affected by the 1990 parental leave reform the findings seem to be sensitive to the exclusion of children born in the pre-reform month June 1990. Hence, although the results suggest a significantly positive and causal reform effect for children of mothers with higher labour force participation rates one has to be cautious about drawing any final policy conclusions. Nevertheless, as already highlighted and demonstrated by Carneiro, Løken and Salvanes (2010) for Norway it seems to be very important for the analysis to differentiate between mothers who are

eligible or ineligible for parental leave. Furthermore, the estimated reform effects are only informative about the *net impact* of the parental leave extension on cognitive child outcomes at age 15 (reduced form estimations). This is true for the present as well as for the majority of the cited evaluation studies (which all lack relevant information on eligibility, actual leave taking and maternal time investments in child rearing). Hence, it remains an important task for future research to investigate in more depth the transmission channels through which parental leave mandates might affect child outcomes.

9 References

Almond, Douglas and Janet Currie. 2010.

15827 [published in David Card and Orley Ashenfelter (eds.), Handbook of Labor Economics (2011), Vol. 4b, Chapter 15: 1315-1486.]

AMS. Public Employment Service Austria (Arbeitsmarktservice Österreich), Labour Market Statistics Database. Available at http://iambweb.ams.or.at/

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. ation of *Journal of the American Statistical Association*, 91 (434): 444-455.

Baker, Michael and Kevin Milligan. 2008. *Journal of Health Economics*, 27 (4): 871-887.

Baker, Michael and Kevin Milligan. 2010a. *Journal of Human Resources,* 45 (1): 1-32.

Baker, Michael and Kevin Milligan. 2010b.

Annual Meeting of the American Economic Association, January 6-8, 2011 (version from December 2010).

Baum II, Charles L. 2003. rly Maternal Employment Harm Child *Journal of Labor Economics*, 21 (2): 409-448.

Bedard, K., and E. Dhuey.

International evidence of long- *The Quarterly Journal of Economics* 121 (4): 1437-1472.

Belmont, Lillian, Zena Stein and Patricia Zybert. 1978.

Order: Effect on Intellectual Ability in Two- *Science, New Series*, 202 (4371): 995-996.

Berger, Lawrence M., Jennifer Hill and Jane Waldfogel. 2005. *The Economic Journal*, 115 (501): F29-F47.

Bernal, Raquel. 2008. Child Care on *International Economic Review*, 49 (4): 1173-1209.

Bernal, Raquel and Michael P. Keane. (2011). *Journal of Labor Economics*, 29 (3): 459-512.

Black, Sandra, Paul J. Devereux and Kyell G. Salvanes. 2005. *The Quarterly Journal of Economics*, 120 (2): 669 700.

Black, Sandra, Paul J. Devereux and Kyell G. Salvanes. *The Review of Economics and Statistics*, 93 (2): 455-467.

Blau, Francine D. and Adam J. Grossberg. 1992. Children's Cognitive Deve *The Review of Economics and Statistics*, 74 (3): 474-481.

BMUJF. 1999a. - Band 1, Zur Situation von Federal Ministry of Environment, Youth and Family. [Austrian Family Report 1999 - Vol. 1, On the situation of families and family policy in Austria]; available at: http://www.bmwfj.gv.at/Familie/Familienforschung/Seiten/4Familienbericht1999.aspx

BMUJF. - Band 2, Partnerschaften zur Vereinbarkeit und Neuverteilung von Betreuungs- Federal Ministry of Environment, Youth and Family. [Austrian Family Report 1999 - Vol. 2, Partnerships for reconciliation and redistribution of care and employment activities] available at: http://www.bmwfj.gv.at/Familie/Familienforschung/Seiten/4Familienbericht1999.aspx

Breit, Simone and Claudia Schreiner. 2007. -
[Sampling design and sample] in Claudia Schreiner and Günter Haider (Eds.): PISA 2006. Internationaler Vergleich von Schülerleistungen. Technischer Bericht. [PISA 2006. International comparison of student achievement. Technical Report.] Austria.

Brooks-Gunn, Jeanne, Wen-Jui Han and Jane Waldfogel. 2002.
Employment and Child Cognitive Outcomes in the First Three Years of Life:
Child Development, 73 (4): 1052-1072.

BSMG. 2003.
Sicherheit,
Generationen und Konsumentenschutz. Wien, Österreich.

Buckles, Kasey S. and Elizabeth L. Munnich. 2011.
Association of America 2011 (PAA) in Washington, DC.

Carneiro, Pedro, Katrine Løken and Kjell G. Salvanes. 2010. A flying start? Long Term Consequences of Maternal Time Investments in Children During Their

Currie, Janet. 2003.
Children and Families conference on Work, Family, Health and Well-Being, Washington D.C. June 16-18, 2003, and published in Suzanne Bianchi and Lynn Casper (eds.) Work, Family, Health and Well-Being (Mahwah NJ: Lawrence Earlbaum Associates Inc.) 2005. Available at: http://www.econ.columbia.edu/currie/Papers/When_do_we_know.pdf

Datta Gupta, Nabanita and Marianne Simonsen. 2010. -cognitive child
Journal of Public Economics,
94 (1-2): 30-43.

Dörfler, Sonja. 2004. - Status Quo und
-family Childcare in Austria Status Quo and Demand], Austrian Institute for Family Studies (ÖIF), Working Paper No. 43.

Dustmann, Christian and Uta Schönberg. 2010.

-

paper version from October 2010.

European Community. 1996.

framework agreement on parental leave concluded by UNICE, CEEP and the

-

lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0034:EN:NOT

Fuchshuber, Eva. 2006. rs Die Repräsentanz von Frauen in Führungspositionen in österreichischen Unternehmen sowie in der

The Representation of Women in Leading Positions in Austrian Enterprises and Public Administration], Bundesministerium für Gesundheit und Frauen, Sektion II. Vienna. March 2006.

Gennetian, Lisa A., Heather D. Hill, Andrew S. London and Leonard M. Lopoo. 2010. Maternal employment and the health of low-

Journal of Health Economics, 29 (3): 353-363.

Gregg, Paul, Elizabeth Washbrook, Carol Propper and Simon Burgess. 2005.

The Economic Journal, 115 (501): F48-F80.

Grossman, Michael. outcomes. in Hanushek, Eric

Chapter 10: 577-633.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. 2008.

Science, 320 (5880): 1164-1165.

Han, Wen-Jui, Christopher Ruhm and Jane Waldfogel. 2009.

- *Journal of Policy Analysis and Management*, 28 (1): 29-54.

Hansen, Kristine and Denise Hawkes. 2009. Early Childcare and Child Development. *Journal of Social Policy*, 38 (2): 211-239.

Hill, Jennifer L., Jane Waldfogel, Jeanne Brooks-Gunn, and Wen-Jui Han. 2005.

Developmental Psychology, 41(6): 833-850.

Hoem, Jan M., Alexia Prskawetz, and Gerda Neyer. 2001a.
Conservative Adjustment? The Effect of Public Policies and Educational
Attainment on Third Births in Austria, 1975-
249 261.

Hoem, Jan M., Alexia Prskawetz, and Gerda Neyer. 2001b.
Conservative Adjustment? The Effect of Public Policies and Educational
A
Institute for Demographic Research, WP 2001-016. Rostock:
http://www.demogr.mpg.de/Papers/Working/wp-2001-016.pdf

Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at www.humanfertility.org [data downloaded on 17/09/2010].

Keil, Achim. 2005.
rt of compulsory schooling: a
-site of the School
Supervisory Association of the Federal Republic of Germany (Konferenz der Schulaufsicht in der Bundesrepublik Deutschland, KSD e.V.). Available at: http://ksdev.de/Schulpflicht.htm [downloaded on 23/03/2011]

Kreimer, Margareta. 2002.
http://elliscambor.mur.at/pdf/vaeterkarenz.pdf

Lalive, Rafael, Analía Schlosser, Andreas Steinhauer and Josef Zweimüller. 2010.
Meeting of the American Economic Association, January 6-8, 2011 (version as of December 31, 2010).

Lalive, Rafael and Josef Zweimüller. 2009.
The Quarterly Journal of Economics, 124 (3): 1363 1402.

Liu, Qian and Oskar Nordstrom Skans. 2010. The Duration of Paid Parental Leave
The B.E. Journal of Economic Analysis & Policy, 10(1) (Contributions), Article 3.

Machin, Stephen and Tuomas Pekkarinen. 2008. lobal Sex Differences in Test *Science*, 322 (5906): 1331-1332.

Mammen, Kristin. 2011. *Journal of Population Economics,* 24 (3): 839-871.

Minnesota Population Center. 2011. Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota. Original data provided by the National Bureau of Statistics, Austria.

Morrill, Melinda S. 2011. nt on the health of school- *Journal of Health Economics*, 30 (2): 240-357.

Mühlenweg, Andrea M. effects within grades on victimization in elementary school*." Economics Letters,* 109 (3): 157-160.

Neuwirth, Norbert. 2004. - An Estimation *Austrian Institute for Family Studies (ÖIF) Working Paper, No. 46.*

Neyer, Gerda R. 2003. *Journal of Population and Social Security: Population Study,* Supplement to Volume 1: 46-93.

OECD. 1989. -

OECD. 2009a. http://www.pisa.oecd.org/dataoecd/0/47/42025182.pdf

OECD. 2009b. Available at: http://browse.oecdbookshop.org/oecd/pdfs/browseit/9809031E.PDF

OECD. Available at: www.oecd.org/els/social/family/database

Pettersson-Lidbom, Per, and Peter Skogman Thoursie. 2009. Labour Market Policy Evaluation (IFAU), Working Paper No. 2009:7.

Ruhm, Christopher J. 1998.

The Quarterly Journal of Economics, 113 (1): 285-317.

Ruhm, Christopher J. 2000. *Journal of Health Economics*, 19 (6): 931-960.

Ruhm, Christopher J. 2004.

Journal of Human Resources, 39 (1): 155-192.

Ruhm, Christopher J. 2008.

Labour Economics, 15 (5): 958-983.

Schneeweis, Nicole and Rudolf Winter-Ebmer. 2007.

Empirical Economics, 32 (2-3): 387-409.

Schneeweis, Nicole and Martina Zweimüller. 2009.

the Analysis of the Welfare State, Working Paper No. 0920.

Schreiner, Claudia, Simone Breit, Ursula Schwantner and Andrea Grafendorfer. 2007. PISA 2006. Internationaler Vergleich von Schülerleistungen. Die Studie

Achievements. An Overview of the Study] Graz: Leykam. Austria.

Sprietsma, Maresa. rst grade of primary school on long-term scholastic results: international comparative evidence using PISA *Education Economics*, 18 (1): 1-32.

2009.

Second and Third-Birth Rates

Working Paper No. 7/2009.

Stanzel-Tischler, Elisabeth and Simone Breit.

(Hrsg.): Nationaler Bildungsbericht Österreich 2009. Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen. Graz: Leykam, 15-32.

Statistik Austria.

statistics 2009/2010], Wien: Verlag Österreich GmbH.

Tanaka, Sakiko. 2005.

The Economic Journal, 115 (501): F7-F28.

Waldfogel, Jane, Wen-Jui Han and Jeanne Brooks-Gunn. 2002.

Demography,

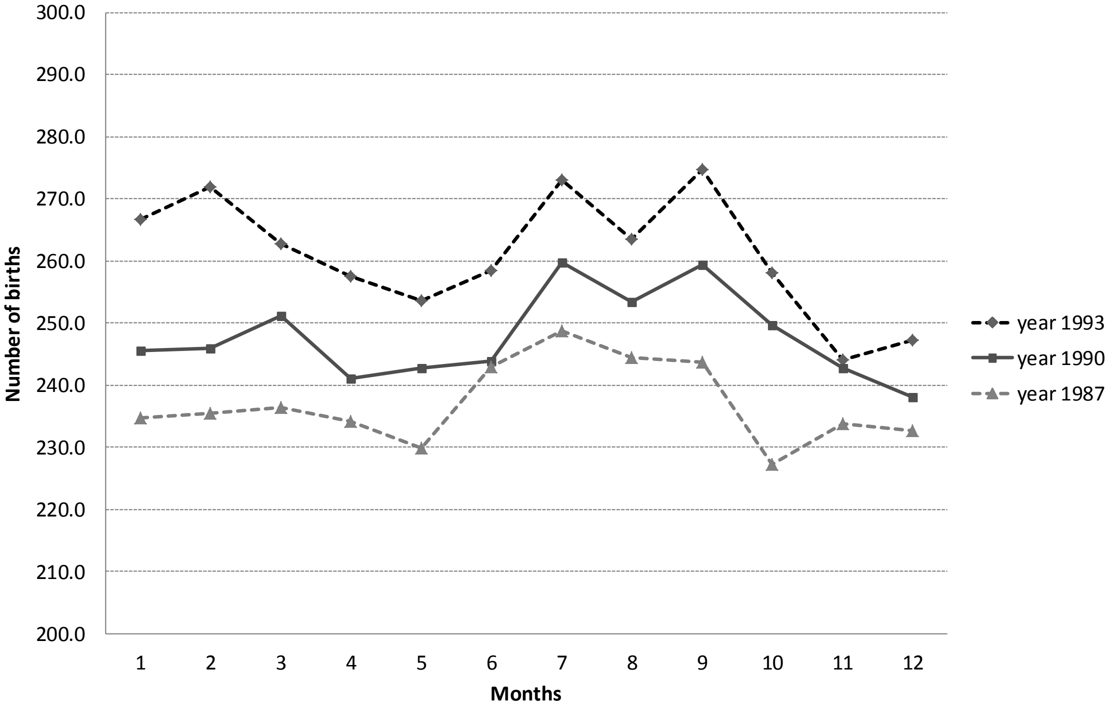
39 (2): 369-392.

World Bank. 2011. Gender Statistics Database. Available at http://data.worldbank.org/data-catalog/gender-statistics [data downloaded on 02/05/2011].

Würtz Rasmussen, Astrid. -related leave: The effect on children's long- *Labour Economics*, 17 (1): 91-100.

Appendix

Figure A 1: Seasonal birth pattern (average number of live births per day across months)



Notes: Birth data refer to the resident population of Austria (permanent residents), irrespective of citizenship, and do not include births registered abroad. Migrants having stayed in Austria less than 3 months are not counted in the resident population. Source: Human Fertility Database (monthly birth numbers). Daily number of births based on own calculations.

Table A 1: Female labour force participation rates in selected countries (% of female population ages 15-64)

	1980	1985	1990	1995	2000	2005	2009
Australia	52.0	54.4	61.9	64.2	65.5	68.2	70.1
Austria	49.9	50.8	55.3	62.2	62.3	65.5	68.4
Denmark	71.7	74.8	77.6	73.3	75.4	75.8	76.5
France	55.2	56.5	57.7	60.5	62.3	64.8	65.5
Germany	52.0	52.3	57.8	61.5	63.5	68.0	71.2
Italy	39.7	39.9	43.6	42.5	46.2	50.4	51.8
Netherlands	48.2	48.4	52.4	58.2	65.6	70.0	73.7
Norway	61.8	67.1	69.9	72.2	76.2	74.5	75.7
Spain	33.1	34.5	41.4	45.8	51.8	58.2	63.1
Sweden	75.0	78.9	81.9	77.3	74.8	76.4	77.0
Switzerland	64.9	66.0	68.2	69.1	71.7	74.3	76.4
United Kingdom	56.3	61.0	66.1	65.9	67.7	68.7	69.4
United States	59.8	63.9	67.5	69.4	70.4	68.6	68.1

Notes: Labour force participation rate is defined as the proportion of the economically active population aged 15-64. Source: World Bank (2011). Gender Statistics Database.

Table A 2: DID estimates, further specifications

	(A) Additional controls for school programme Two months window (May-Aug)			(B) One month window (Jun-Jul) Main specification (standard control variables)		
	(1) Full sample	(2) High LFP mothers	(3) Low LFP mothers	(4) Full sample	(5) High LFP mothers	(6) Low LFP mothers
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	5.802	20.335***	-1.843	15.827	46.545***	3.851
	(4.643)	(6.742)	(6.559)	(11.000)	(17.205)	(14.990)
Reading	-0.463	24.799**	-13.269*	11.118	55.814***	-8.905
	(6.264)	(10.191)	(6.643)	(11.624)	(18.600)	(14.871)
Science	5.841	26.334***	-5.070	11.370	45.131***	-3.056
	(5.559)	(7.816)	(6.793)	(11.633)	(15.713)	(15.429)
	MALES			**MALES**		
Mathematics	6.482	18.951**	-1.565	22.142*	59.477***	3.493
	(6.176)	(8.810)	(9.276)	(12.254)	(13.847)	(18.967)
Reading	-1.770	32.890**	-19.939*	22.062	89.072***	-7.507
	(8.348)	(14.420)	(9.674)	(15.176)	(24.605)	(21.125)
Science	4.479	41.727***	-16.264	19.684	80.092***	-8.818
	(7.316)	(9.158)	(10.110)	(15.128)	(17.602)	(22.173)
	FEMALES			**FEMALES**		
Mathematics	6.005	17.721	-1.131	8.832	29.720	1.123
	(6.815)	(12.052)	(9.133)	(17.555)	(32.105)	(24.157)
Reading	0.230	15.905	-6.717	0.361	25.030	-8.991
	(9.373)	(16.923)	(8.592)	(16.869)	(25.715)	(22.423)
Science	7.763	7.868	7.344	3.288	6.676	3.341
	(8.234)	(12.725)	(9.874)	(17.108)	(23.328)	(24.125)
Controls						
Parental background	✓	✓	✓	✓	✓	✓
School programme	✓	✓	✓	-	-	-

Notes: The upper panel includes only male, the lower panel only female students. All regressions include dummy variable controls for survey year, birth months and for all children born post June. The control variab
attainment, school location, migration background and for the five different school types. The sample size for the pooled samples (top panel; row 1 and 3) are 2,840 and 1,386 respectively. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table A 3: DID estimates, including only one post-reform month (July)

	High LFP mothers			Low LFP mothers		
	(1) Mar-Jul	(2) Apr-Jul	(3) May-Jul	(4) Mar-Jul	(5) Apr-Jul	(6) May-Jul
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	8.229	12.587	18.436	1.140	-0.269	-6.783
	(15.021)	(15.374)	(16.278)	(9.949)	(10.906)	(13.139)
Reading	17.776	22.329	28.804*	-7.819	-9.188	-16.252
	(13.783)	(14.203)	(15.804)	(10.120)	(10.705)	(12.468)
Science	15.428	20.104	26.311*	-5.060	-6.768	-10.773
	(12.239)	(12.719)	(13.670)	(10.978)	(11.397)	(13.389)
Observations	1,145	916	681	2,346	1,883	1,413
	MALES			**MALES**		
Mathematics	20.046	20.516	22.530	-2.083	-2.593	-8.195
	(13.042)	(13.605)	(15.253)	(10.263)	(11.900)	(14.752)
Reading	45.181***	47.766**	52.335**	-9.156	-12.951	-18.941
	(16.062)	(17.502)	(19.672)	(11.375)	(12.760)	(15.171)
Science	41.839***	44.333***	49.019***	-19.719*	-21.487*	-25.685
	(12.292)	(13.026)	(14.065)	(11.359)	(12.091)	(15.982)
Observations	581	457	345	1,200	950	720
	FEMALES			**FEMALES**		
Mathematics	-3.534	1.535	10.342	5.569	3.115	-3.839
	(28.181)	(29.468)	(31.116)	(15.939)	(17.161)	(20.301)
Reading	-11.181	-3.387	10.113	-0.699	-2.565	-15.003
	(22.966)	(23.740)	(26.898)	(16.582)	(17.788)	(21.782)
Science	-11.819	-7.059	0.125	12.778	10.421	6.205
	(20.282)	(21.443)	(23.573)	(18.620)	(19.504)	(21.794)
Observations	564	459	336	1,146	933	693

Notes: The reported estimated treatment effects stem from separate estimations of different specifications based on the Austrian PISA data 2006 and 2003. All regressions include dummy variables for month of birth, a year dummy for 2006, a dummy variable for all children born after June. The control variables on

migration background of the family. Robust standard errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table A 4: DID estimations excluding observations from Vienna.

	High LFP mothers			Low LFP mothers		
	(1) Mar-Oct	(2) Apr-Sep	(3) May-Aug	(4) Mar-Oct	(5) Apr-Sep	(6) May-Aug
		MALES			**MALES**	
Mathematics	15.760	17.056	15.945	3.062	6.102	1.903
	(12.834)	(11.867)	(15.894)	(9.063)	(9.876)	(10.871)
Reading	32.536**	36.820**	42.002**	-12.455	-11.839	-15.642
	(13.646)	(14.137)	(16.713)	(10.876)	(12.445)	(11.604)
Science	35.190***	40.442***	44.329***	-9.637	-7.091	-9.195
	(12.157)	(12.663)	(14.160)	(9.581)	(9.339)	(10.815)
Observations	784	586	391	1,590	1,205	824
		FEMALES			**FEMALES**	
Mathematics	6.716	11.041	24.195	2.743	0.766	-2.255
	(10.694)	(11.801)	(16.923)	(9.447)	(10.843)	(14.919)
Reading	-1.750	3.669	20.118	5.595	2.452	0.509
	(15.200)	(15.160)	(20.835)	(9.983)	(11.019)	(15.014)
Science	-4.031	-0.906	9.784	8.535	5.756	6.486
	(12.914)	(13.848)	(17.708)	(10.845)	(12.040)	(14.731)
Observations	785	602	396	1,695	1,273	845

Notes: The reported estimated treatment effects stem from separate estimations of different specifications based on the Austrian PISA data 2006 and 2003. The sample excludes children living in Vienna. All regressions include dummy variables for month of birth, a year dummy for 2006, a dummy variable for all children born after June. The control variables on parental background include dummy variables for
obust standard
errors in parentheses (clustered by school programme, school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table A 5: DDD estimates, including only two post-reform months (July and August)

	High LFP mothers			Low LFP mothers		
	(1) Mar-Aug	(2) Apr-Aug	(3) May-Aug	(4) Mar-Aug	(5) Apr-Aug	(6) May-Aug
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	18.991*	21.884*	22.765	-5.564	-9.276	-17.131
	(9.963)	(11.259)	(14.804)	(10.109)	(11.773)	(14.101)
Reading	18.470	20.374	25.079	-22.205*	-25.883*	-30.991*
	(13.437)	(15.170)	(17.315)	(12.420)	(13.402)	(15.499)
Science	28.977**	33.443**	37.808**	-14.057	-16.326	-20.919
	(11.869)	(13.328)	(15.267)	(11.478)	(13.253)	(14.809)
Observations	3,143	2,645	2,158	4,877	4,111	3,325
	MALES			**MALES**		
Mathematics	28.337**	32.937**	38.109**	-32.050*	-33.276*	-37.219**
	(13.109)	(13.506)	(18.223)	(16.075)	(17.255)	(16.667)
Reading	32.141*	36.247*	40.010*	-53.978**	-55.936**	-57.088**
	(16.556)	(18.176)	(22.573)	(19.885)	(21.458)	(22.137)
Science	48.457***	56.952***	66.932***	-50.191***	-52.185**	-53.071***
	(10.736)	(11.991)	(15.812)	(16.658)	(18.681)	(18.600)
Observations	1,615	1,350	1,113	2,410	2,014	1,636
	FEMALES			**FEMALES**		
Mathematics	11.277	11.569	9.412	16.768	11.147	2.579
	(16.250)	(18.794)	(23.564)	(11.624)	(15.431)	(21.739)
Reading	1.167	3.640	13.632	8.589	2.491	-8.546
	(18.502)	(21.409)	(21.205)	(12.821)	(16.398)	(23.939)
Science	9.526	9.598	10.392	17.281	14.894	9.852
	(19.197)	(21.652)	(23.252)	(13.708)	(17.689)	(22.128)
Observations	1,528	1,295	1,045	2,467	2,097	1,689

Notes: The reported estimated treatment effects stem from separate estimations of different specifications. All regressions include dummy variables for month of birth, year and country fixed effects, a dummy variable for all children born after June, interaction effects between year and the post June dummy, year and country, country and post June . The control variables on parental background include dummy tional attainment, school location, and migration background of the family. Robust standard errors in parentheses (clustered by school track (more/less academic), school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.

Table A 6: DDD estimates, including only one post-reform month (July)

	High LFP mothers			Low LFP mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mar-Jul	Apr-Jul	May-Jul	Mar-Jul	Apr-Jul	May-Jul
	MALES + FEMALES			**MALES + FEMALES**		
Mathematics	23.787*	27.064*	28.006	-12.450	-16.021	-24.164
	(13.928)	(14.826)	(17.540)	(12.830)	(14.956)	(17.880)
Reading	29.370*	31.610*	35.799*	-20.369	-23.877	-29.145
	(16.311)	(17.557)	(18.973)	(14.704)	(16.108)	(18.571)
Science	36.203**	41.148**	45.302***	-19.369	-21.599	-26.645
	(14.061)	(15.300)	(16.807)	(14.772)	(16.903)	(19.038)
Observations	2,566	2,068	1,581	4,017	3,251	2,465
	MALES			**MALES**		
Mathematics	26.028	30.699	35.408	-37.074***	-37.652**	-41.197***
	(21.021)	(20.366)	(23.135)	(12.042)	(14.257)	(14.221)
Reading	43.741*	48.882*	52.053*	-51.579***	-52.982***	-53.417***
	(22.900)	(24.392)	(27.165)	(16.004)	(17.679)	(18.416)
Science	54.234***	63.334***	73.115***	-57.985***	-59.688***	-60.479***
	(18.995)	(18.917)	(20.420)	(10.752)	(13.951)	(15.451)
Observations	1,306	1,041	804	2,002	1,606	1,228
	FEMALES			**FEMALES**		
Mathematics	21.529	22.319	21.237	10.292	4.967	-4.348
	(21.740)	(24.856)	(30.430)	(20.344)	(23.908)	(30.504)
Reading	6.228	8.959	18.564	3.392	-2.553	-14.082
	(18.633)	(19.764)	(21.419)	(23.260)	(27.553)	(34.500)
Science	16.580	17.040	18.002	15.769	13.669	7.646
	(19.553)	(22.626)	(26.049)	(24.241)	(28.202)	(32.892)
Observations	1,260	1,027	777	2,015	1,645	1,237

Notes: The reported estimated treatment effects stem from separate estimations of different specifications. All regressions include dummy variables for month of birth, year and country fixed effects, a dummy variable for all children born after June, interaction effects between year and the post June dummy, year and country, country and post June . The control variables on parental background include dummy

Robust standard errors in parentheses (clustered by school track (more/less academic), school location, and gender). Estimations weighted by individual inverse probability weights provided in the PISA data set. *** p<0.01, ** p<0.05, * p<0.1. Source: PISA data set (OECD), own calculations.