The Long Term Effect on Children of Increasing the Length of Parents' Birth Related Leave*

Astrid Würtz
Dept. of Economics, Aarhus School of Business, University of Aarhus
awu@asb.dk

January, 2007

PRELIMINARY, please do not quote without permission

Abstract

The length of parents' total birth related leave was increased with almost 50% in 1984 in Denmark. The previous length of the Danish maternity leave was long - also compared to e.g. the U.S. today. This paper investigates the long term effects on children of increasing the length of birth related leave from 14 to 20 weeks. We can identify the causal effect of the leave reform and it will be estimated whether such a large increase in the leave period has a large measurable and persistent effect on children's cognitive and educational outcomes. A 100% sample of the Danish population born in June and July 1983, 1984, and 1985 and a dataset with the Danish PISA-2000 scores are used for the estimations. Preliminary results suggest that there is not a positive effect on children's cognitive outcomes from increasing the birth related leave period from 14 to 20 weeks.

JEL Classification: J13, J18, D13
Keywords: Maternity leave, parental leave, child development

*This paper has partly been written while visiting Dept. of Economics, University of Washington, Seattle. I would like to thank Shelly Lundberg for suggestions to the paper and further gratefully acknowledge comments and suggestions from participants at the "Summer Tan Bag Seminar" at University of Washington, especially comments from Elaina Rose and Levis Kochin. Furthermore, I would like to thank Nina Smith and Nabanita Datta Gupta for helpful suggestions.

1 Introduction

The length of parents' birth related leave was increased with almost 50% in 1984 in Denmark. The previous length of the Danish maternity leave was quite long, also compared to e.g. the U.S. today. So, what is the effect of adding 6 weeks of parental leave to the existing 14 weeks of maternity leave increasing the length of the total birth related leave with almost 50%? Is there a persistent effect on the children when the maternity leave is already 14 weeks to start with? This paper investigates the long term effects on children of this policy reform in Denmark in 1984. Children affected by the reform are born in 1984 and they are therefore 21 years old by the end of 2005. Having data from 1984 to 2005 we can address the reform's effect on high school enrollment, high school completion, and high school grade point averages (GPAs). For some of the children we also use PISA reading scores from OECD's PISA-2000 study as the outcome measure. The reading abilities are tested when the children are 15 years old.

Women in Denmark are actively participating in the labor market, and as shown in Figure 1 the labor force participation for Danish women were high already in the 1980s, especially compared to e.g. American women. Therefore, access to daycare and maternity leave is very important for Danish women and has been throughout the period we investigate.

Previous studies have mostly focused on finding the effect on children of maternal employment and not of maternity or parental leave. Estimating the effect on children caused by maternal employment may not be as reliable as estimating the effect of birth related leave length because there is a selection issue related to maternal employment. Not all women choose to work but almost all women who give birth choose to take birth related leave in Denmark. Therefore, there will not be a selection problem in relation to investigating the effect on children of mothers' birth related leave.

The methods used in this paper to identify the causal effect of the leave reform is regression discontinuity (RD) design and differences-in-differences. It is investigated whether an increase in the length of birth related leave is beneficial for children, i.e. has a measurable, positive impact on their long term cognitive outcomes. Furthermore, we test whether the reform has an impact on enrollment in secondary education. It might be the case that an increase in the total birth related leave will not have a (long term) impact on children because the maternity leave is already long (more than 3 months) before the reform. This is an empirical question which is answered using a full sample of Danish children born in the months around the entry into force of the policy reform. In addition, a dataset with Danish PISA-2000

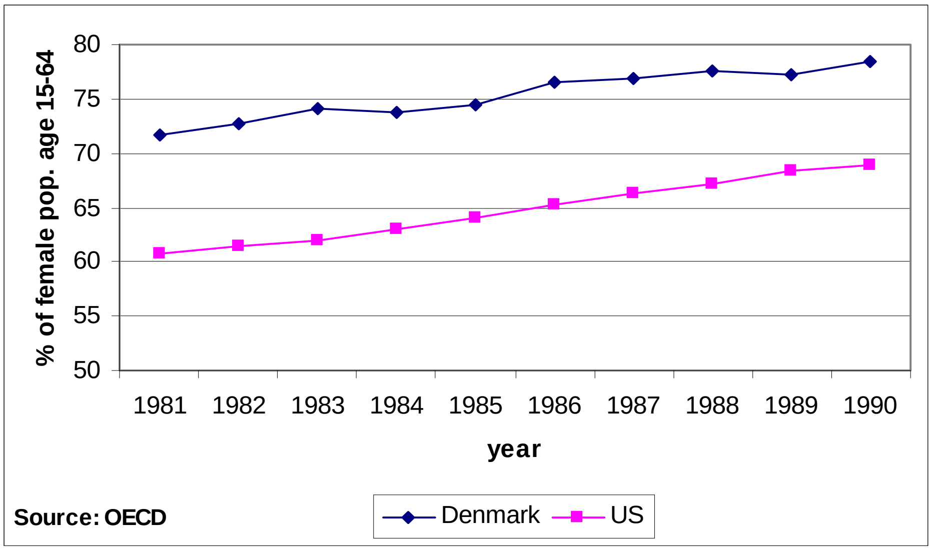

Figure 1: Female labor force participation in Denmark and the US in 1981 to 1990.

scores is also used. These datasets are thoroughly introduced in section 5.

The contribution of this paper is threefold. First, using the method of RD to estimate the effect of the length of parents' birth related leave on children's cognitive and educational outcomes has not been done earlier. Second, long term effects of the length of maternity and parental leave are investigated. Usually, only shorter term outcomes are estimated due to data limitations. Third, policy suggestions for countries that do not have a long maternity or parental leave today can be provided from this study (given the institutional settings are fairly similar to those in Denmark in the 1980s).

In section 2 a short literature review of studies relevant for this paper is presented. In section 3 leave reforms during the 1980s are briefly summarized, and in section 4 the method of regression discontinuity is explained. Section 5 gives an introduction to the datasets used in the estimations, and the empirical results are presented in section 6. Finally, section 7 concludes.

2 Literature Review

Studies on maternity and parental leave focus mostly on the effect of the leave period on either maternal health, child health, maternal employment and wages, or fertility. A selection of studies concerned with these issues are presented in this section along with some important studies about or using the method of regression discontinuity.

A study using data from the U.S. investigates whether the length of maternity leave affects maternal health, see Chatterji and Markowitz (2005). They find that longer maternity leave reduces cases of depression among working mothers, i.e. there is a positive effect of a longer maternity leave. Other papers as e.g. Ruhm (2000), Berger et al. (2005), and Baker and Milligan (2005) have focused on the relationship between child health and maternity leave. Ruhm (2000) and Berger et al. (2005) find that there is a positive relationship between child health and leave duration, and more generous paid leave reduces deaths of infants and young children. Baker and Milligan (2005) on the other hand do not find evidence of an effect of time at home on infant health using Canadian data.

Baker and Milligan (2005) also look at the effect of maternity leave on women's labour supply. They find that 17-18 weeks of maternity leave do not increase the time mothers spend at home whereas leave lengths up to 70 weeks seem to further increase the time spent at home. Pylkkänen and Smith (2004) look at Danish and Swedish mothers' career interuptions due to childbirth. Their study shows that family friendly policies have a bigger impact on Swedish mothers but that both Danish and Swedish mothers' behaviour are affected by the economic incentives and leave periods of the fathers. Finally, Lalive and Zweimüller (2005) investigate the effect of a 1990 Austrian policy reform. They study the effect of parental leave on fertility and women's return-to-work in the private sector and find a big positive effect on fertility of a longer parental leave. It simply pays to get several children shortly after each other. Longer leave also increases women's time off work.

Many studies focus on the effect on children of maternal employment. There is an extensive literature on this subject but the results vary a lot. Baum (2003) finds that in the U.S. mothers' return to work within the first three months of the child's life result in lower cognitive test scores for the children. Ruhm (2004) also uses U.S. data and finds a negative effect on children from early maternal employment, i.e. maternal employment in the first and the three first years after birth. Ruhm (2005) finds that maternal employment is bad for advantaged children but not bad (or even good) for disadvantaged children. Both cognitive outcomes and health (weigth) are considered in this study. Gregg et al. (2005) study data from the U.K. and finds that the type of child care seems to be an important determinant of children's outcome along with maternal employment. The effect of early maternal employment on children's cognitive development between ages 4 and 7 is investigated. Formal child care substitution is better for children than informal child care arrangements according to this study. From this selection of studies it is clear that the effect on children from maternal employment ranges from negative to neutral or even positive and the mixed results may be caused by selection issues in relation to the group of employed mothers.

The literature on regression discontinuity started with the study by Thistlethwaite and Campbell (1960) and has developed since then. The most famous paper within this literature is probably the study by Angrist and Lavy (1999) where Maimonides' rule is used to estimate the effect of class size on scholastic achievement. They find that lower class size increases test scores for children in 4th and 5th grade but not in 3rd grade. An important contribution to the regression discontinuity literature is Hahn et al. (2001). They define and introduce sharp and fuzzy design and clarify which assumptions are needed for identification in a RD design. Finally, the study most similar to the study in this paper is Lalive and Zweimüller (2005). As mentioned above, they use an Austrian leave reform to investigate the effect of parental leave on maternal employment and fertility and they do that in a RD design. They focus only on mothers working in the private sector, though. None of the studies in the regression discontinuity literature investigate the effect of a policy reform on children's long term outcomes. The use of RD design to estimate the effect of the length of parents' birth related leave on children, estimating on a full sample of children born in the relevant period, and at the same time focusing on long term outcomes are the main contributions of this study.

3 Leave Reforms in Denmark in the 1980s

For a very long time period women in Denmark have been entitled to maternity leave. Focusing only at the 1980s we see that in 1983 mothers are allowed to take 14 weeks of maternity leave with income dependent compensation, max. 2008 dkr. (about $335) per week. The local municipality pays the compensation. In addition to the leave after birth, mothers are also entitled to 4 weeks of leave before child birth. Fathers are not entitled to any birth related leave in 1983, i.e. the 44 fathers that take birth related leave pay for it themselves.

A policy reform is implemented July 1st, 1984. The reform extends the leave period to 20 weeks after child birth but the last 6 weeks are parental leave so fathers can get compensation in up to 6 weeks (after week 14). Only one of the parents can get compensation at one point in time. Fathers are furthermore also guaranteed 2 weeks of leave with compensation right after the child's birth. The reform implies that children born in 1984 will not necessarily all have had the same amount of "home time" with their mothers (or fathers)[1], so children in the same school class are different with respect to this[2]. The policy change therefore created a "Natural Experiment".

Another similar reform is implemented July 1st, 1985. The total leave period after birth is extended to 24 weeks where the last 10 weeks are parental leave. Fathers can therefore take the last 10 weeks of leave instead of the mother.

Table 1. Average weeks of leave for men and women in Denmark in 1983 and 1985.

1983	**Men**	**Women**
Avr. weeks of leave	6.12	15.06

1985 Men*

Weeks after birth	Jan-Jun	Weeks after birth	Jul-Dec
1-2	1.8	1-2	1.9
15-20	3.4	15-24	4.5
21-	3.5	25-	4.6

1985 Women**

Avr. weeks of leave	Jan-Jun	Jul-Dec
Total	24.5	29.4
After birth	19.3	23.2

* Excluding self-employed men in 1985. 95% of the men took less than 3 weeks of leave in 1985.

** Excluding self-employed women in 1985.

Source: Statistics Denmark: Statistiske Efterretninger, 1985 and 1987.

As shown in Table 1 mothers on average had 15.06 weeks of leave in

[1] In practice most of the leave is taken by mothers.

[2] School starts in August in Denmark and all children 6 years old in the current year can start in school. This means that in a school class almost all children are born in the same year and there can be up to almost a year in age difference between the oldest and the youngest children. A few children of course start too early or too late in school but the majority of children in a school class are born in the same year.

1983, i.e. the leave before child birth is also recorded in this number. The 44 fathers that took leave and payed for it themselves on average took 6.12 weeks. The numbers for average weeks of leave before and after July 1st in 1984 are unfortunately not available from Statistics Denmark so instead we look at 1985. As mentioned, there was a similar reform to the 1984 reform, and we see that before July 1st, 1985 almost all women took the 20 weeks of leave they were entitled to after birth and after July 1st almost all of them took the 24 weeks they were entitled to. Following this, we believe that the same pattern holds for mothers of children born in 1984 since the available amount of leave in 1984 is lower than in 1985. This will of course be checked directly in the data when it arrives.

4 The Method of Regression Discontinuity

The method of regression discontinuity is a useful method for determining whether a program or treatment is effective. The classical problem is that individuals have either been treated or not treated but cannot be in both categories at the same time. Therefore, it is not possible to use the optimal strategy which is to compare the effect of treatment and no treatment for the exact same individual.

The idea of RD is to determine the effect of a treatment, T_i, on an outcome, Y_i, where the treatment assignment function is discontinuous at point B. Intuitively, we compare individuals very close to the discontinuity point so we expect them to be similar, except for the fact that they have been exposed to different treatments. That is, their value of the underlying targeting variable is just below and just above the discontinuity point but apart from that they have experienced identical environments. The average treatment effect is therefore estimated by comparing average outcome values of those individuals just above and just below B and the treatment effect is identified exactly at B. To ensure identification, a sample ”close” to the discontinuity point must be used and we use 1 month on either side of the discontinuity point as in Lalive and Zweimüller (2005).

What is unique in the RD design is the way individuals are allocated to different groups based solely on a cutoff criterion. Therefore, individuals in different groups do not have to be identical given "pre-program" indicators as in e.g. a randomized experiment. It can for example be the case that individuals are allocated to different groups based on their health or based on a test score. It is assumed that in the absence of the "program" (policy change) the pre-post relationship would be equivalent for the two groups. In this study the cutoff criterion is determined on the basis of a birth date so here we will actually also expect individuals to be equivalent on pre-program indicators. All individuals born before the cutoff value (July 1st, 1984) are assigned to the "control group" whereas individuals born after the cutoff value are assigned to the "treatment group". Treatment is parents' entitlement to the extra 6 weeks of parental leave in connection to the childbirth.

When RD design is well implemented, as we argue it is in this study, inferences are comparable in internal validity to conclusions from randomized experiments, i.e. the policy change investigated indeed caused a change in the outcome indicators; children's cognitive outcomes and educational attainment. RD can therefore be used to investigate causal hypotheses for the immidiate research context.

There are two types of regression discontinuity designs, sharp design and fuzzy design, see Hahn et al. (2001). Under sharp design, treatment is a discontinuous but deterministic function, $f(\cdot)$, of some variable, b_i, where b_i takes on a continuum of values. If treatment assignment is not a deterministic function of b_i, i.e. there are additional variables unobserved to the researcher that determine assignment to treatment, then we have a so-called fuzzy design. In this study a deterministic treatment assignment function can be set up as

$$T_i = f(b_i) = 1(b_i \geq \mathit{July\ 1st,\ 1984}),$$

where b_i is the date of birth. We therefore have a sharp design with discontinuity point at $B = \mathit{July\ 1st,\ 1984}$. Mothers are predicted to take as much leave as possible, i.e. under treatment - the child is born after the discontinuity point - mothers take 20 weeks of leave, otherwise they take 14 weeks of leave. This is illustrated in Figure 2.

The pre-post relationship is well known and therefore we can correctly model it. This along with the fact that we do not - as we argue below - have any spurious discontinuity in the pre-post relationship at the cutoff point ensures the strength and validity of the RD design. Discontinuity of treatment at B is ensured because Assumption 1 is fulfilled. Assumption 2 ensures that the pre-test distribution is continuous which is crucial for identification of the treatment effect because it ensures that the average treatment effect is similar for individuals with values of b_i close to B. Finally, Assumption 3 generalizes the identification strategy to include heterogenous treatment effects instead of only constant treatment effects.

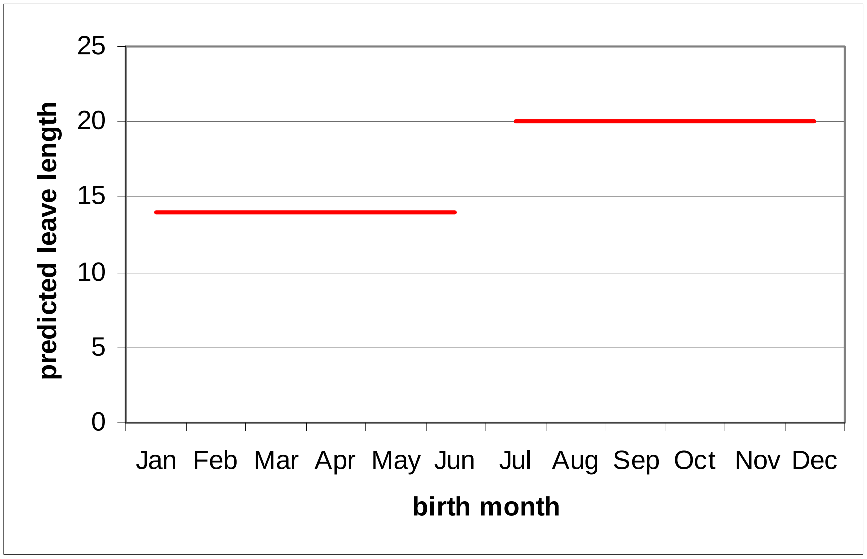

Figure 2: Predicted leave length for mothers based on the month of child birth in 1984.

Assumption 1 The limits

$$T^{+} \equiv \lim_{b \to B^{+}} E\left[T_i | b_i = b\right] \quad \text{and} \quad T^{-} \equiv \lim_{b \to B^{-}} E\left[T_i | b_i = b\right]$$

exist and are *not* equal.

Assumption 2 $E\left[outcome\ absent\ treatment | b_i = b\right]$ is continuous in $b = B$.

Assumption 3 $E\left[treatment\ effect | b_i = b\right]$ regarded as a function of b is continuous at B.

Further, if we assume that T_i is independent of the treatment effect conditional on b_i close to B, then the average treatment effect at B is non-parametrically identified as

$$\frac{y^{+} - y^{-}}{T^{+} - T^{-}},$$

where $y^{+} \equiv \lim_{b \to B^{+}} E\left[y_i | b_i = b\right]$ and $y^{-} \equiv \lim_{b \to B^{-}} E\left[y_i | b_i = b\right]$. Actually, when we have sharp design, $T^{+} - T^{-} = 1$ so the treatment effect is given by the simple expression $y^+ - y^-$, see Hahn et al. (2001). The treatment effect is consistently estimated given consistent estimators of y^+ and y^-. The (weak) conditional independence assumption ensures that individuals do not select into treatment on the basis of their anticipated gains from treatment which is important for the internal validity of the RD design.

If anyone could anticipate the increase in the leave length and change behavior according to that we would have endogeneity problems. But selection on the basis of parents' deliberate choices can be ruled out because of biological limitations. If the reform could be predicted, some parents might plan to have children at a later point in time than what they would otherwise have had. But it might be possible to give birth earlier than after 9 months pregnancy but it is impossible to delay child birth if the birth is natural, i.e. not a caesarean section. Therefore, since we expect parents to prefer having their child born after July 1st instead of before July 1st, we do not have to deal with selection issues with respect to planning or influencing the time of birth. Further, nothing seems to indicate that it was possible to predict the entry into force of the reform more than nine months before July 1st 1984.

The number of births increased in 1984 compared to 1983 as shown in Figure 3. This ended a long time period with decreasing number of births. But is the increase in number of births in 1984 a response to the leave reform? If it is we may have a selected sample. From Figure 4 we see, though, that the increase in number of births in Denmark does not seem to be a response to the reform. The increase in births is evenly spread throughout the year and is not allocated only to the last six months. We would have expected the latter if the increase in fertility was caused by the reform. Figure 4 also shows that there is seasonality in births. There are many births in March to August but fewer births in the first two and last four months of the year. Further, the amount of births in June and July seems to be almost equal in both 1983 and 1984 which makes us even more confident that we do not have sample selection problems. According to Figure 4, even if the policy change could be anticipated individuals did not change behaviour.

For the analytical model to be appropriate it is important that Assumption 1 is fulfilled, i.e. that the cutoff criterion is followed without exception. This means that policy implementation is uniform to all recipients so that they all receive the same entitlement to leave. Trusting the Danish system this definitely seems to be the case with a policy reform based on date of child birth. The probability of receiving treatment changes discontinously as a function of an underlying variable and there is no deviation from the rules.

In this study we have comparison group pretest variance, i.e. there is a

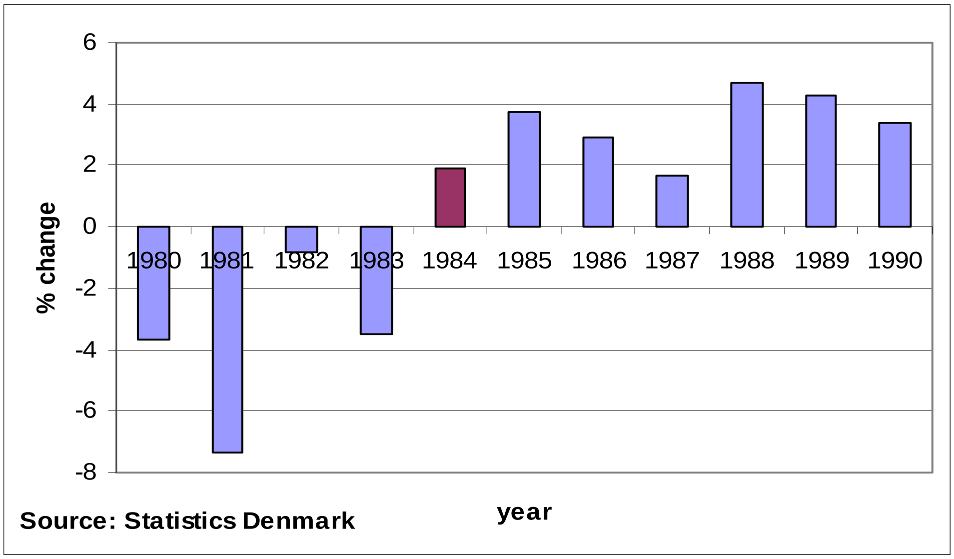

Figure 3: Percentage change in number of births in Denmark in the period 1980 to 1990.

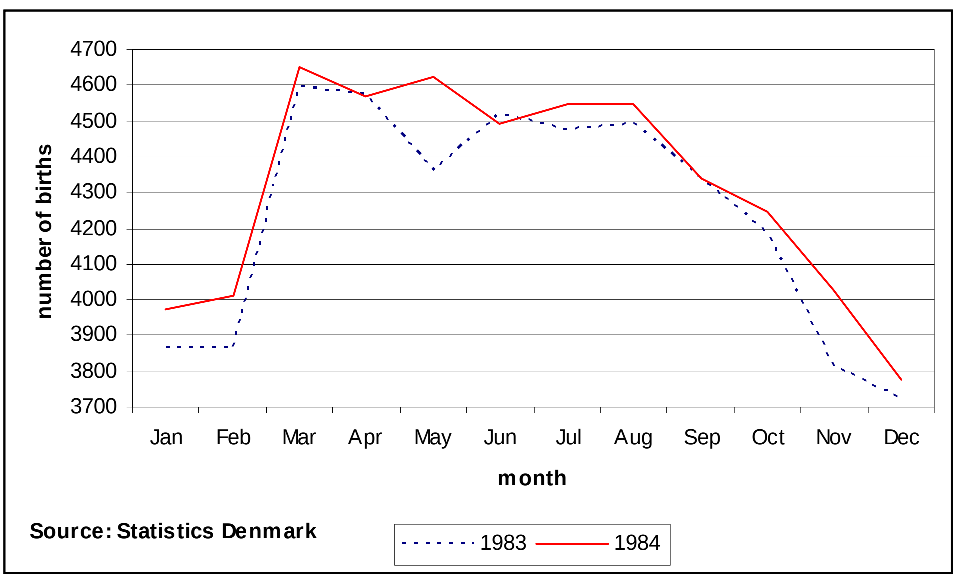

Figure 4: Births per month in Denmark in 1983 and 1984.

sufficient number of pretest values in the comparison group to enable adequate estimation of the true relationship for that group. Our comparison group consists of about 4,500 observations in the 100% sample, but we may have to be more careful with only about 350 observations in the comparison group of the PISA-dataset. Apart from the probability of treatment, individuals on either side of the cutoff point experience almost identical environments. Both groups come from a single continuous pretest distribution with the division between groups determined only by the cutoff. Further, we cannot identify any other major changes in Denmark at this point in time. We therefore do not have to worry about possible spurious discontinuities in the pre-post relationship which happens to coincide with the cutoff point.

Lastly, a more technical issue, regression to the mean is expected but implies no problem in the RD design because it does not impose a discontinuity at the cutoff point. As mentioned, the important thing to worry about in RD design is whether we risk having other discontinuoties at B than the one imposed by the treatment. If we have other discontinuities it is impossible to correctly identify the true effect of the policy change in B.

4.1 Empirical Model Specification

One of the biggest problems when analyzing data using RD design is misspecification of the statistical model[3]. This might lead to biased estimates of the policy effect (or pseudo-effects). To make sure the statistical model is not misspecified one should examine the pre-post relationship visually to determine whether there is any visually discernable discontinuity at the cutoff. A change in the level vertically is the main effect, a change in the slope is an interaction effect, and there can be both. We see in Figure [TBM] that the pre-post distribution is linear and therefore it makes sense to use a RD design for identifying the effect of the policy change.

Visually finding a discontinuity ensures that there must also be one analytically, given there is no selection in the data. Furthermore, visually inspecting the data gives us an idea about the order of the polynomial that fits the data. We then specify higher-order terms and interactions based on the flexion points detected in the visual inspection of the data. For each transformation of the pretest, X, interaction terms are also created by multiplying the polynomial by T. The initial model is finally estimated and refined by removing apparently unnecessary terms. The latter ensures more efficient estimates of the treatment effect.

[3]E.g. if curvilinear data is misspecified into a linear model with a discontinuity at the cutoff point.

Formally, based on the visual inspection in Figure [TBM] we assume that the true function is given as

$$Y_i = B_0 + B_1 X_i + B_2 T_i,$$

where Y is the term for the posttest, X is for the pretest and T is for the dummy coded treatment variable. We can then estimate

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 T_i + \varepsilon_i,$$

where ε_i is the error term. Given that the true model is exactly specified, we get unbiased and efficient estimates of the treatment effect. The parameter of interest is β_2 which measures the main effect of the policy change, i.e. the vertical discontinuity at the cutoff point is estimated by this coefficient.

As mentioned above, we overspecify the model. This ensures us from getting biased estimates, but only at the cost of possibly finding inefficient estimates of the treatment effect. We can always in successive analyses gradually remove higher-order terms if it seems to be necessary. The final (initial) estimation equation is therefore given as

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 T_i + \beta_3 X_i T_i + \beta_4 X_i^2 + \beta_5 X_i^2 T_i + \varepsilon_i.$$

5 Data

Estimations are based on two data sources. First, an administrative register dataset consisting of a 100% sample of the Danish population born in June and July 1983, 1984, and 1985 is used, and secondly, a dataset with the Danish PISA-2000 scores combined with register information from Danish administrative registers is used. The two datasets are presented in the following subsections.

5.1 100% Sample

The total Danish population consists of about 5.4 million people i 2006. In 1984 51,800 children were born, 9035 of these were born in June or July. These children are the focus of this study. In the Danish tax and income registers created by Statistics Denmark these individuals and their parents are followed on a yearly basis from 1984 to 2005 if they have not left the country or died. This enables us to find information on the parents' and children's marital status, residence, education, income, wage, labor market activities, children, etc. Furthermore, we get information on the children's enrollment in education in 2005 and their GPAs from elementary school and high school, if they have completed a high school education. Therefore, we can determine children's long term educational outcome.

In addition to information about children born in June and July 1984 we also have a full sample of children born in June and July 1983 and 1985. We therefore have a dataset that covers the full sample of children born in the relevant time period, i.e. we have the best possible dataset available for the analyses in this study.

[TBW - waiting for the data]

5.2 Danish PISA-2000 sample

PISA is short for the "OECD Programme for International Student Assessment". In the year of 2000 similar tests of 15 year old children were conducted in 32 countries, most of them OECD countries. The PISA-2000 study focused on children's reading abilities but also tested some of the children in mathematics and science. We use the reading score as the outcome measure in this study since all children were tested in reading[4]. Test scores from PISA tests are normalized to an OECD mean of 500 and with a standard deviation of 100. Denmark scored below average in reading with a mean of 497 and a standard deviation of 98. In comparison, Finland scored highest with an average score of 546 in reading and with a standard deviation of 89. For information on the PISA-2000 study, see OECD (2002).

Children participating in the PISA study are equipped with identifiers to combine PISA information with register data. This gives us information on the child and its parents from child birth to the year 2005. Furthermore, PISA-children have answered questionnaires about their family background, home environment etc., and school principals or head administrators on the schools which participated in the PISA test was also asked to respond to a questionnaire with information about the school, teachers etc. This information is linked to the child. We will include many background variables from questionnaires and registers in the estimations to decrease the possible omitted variables bias.

4242 children participated in the Danish PISA test in 2000. Of these, 344 children were born in June and 362 were born in July 1984.

[TBW - waiting for the data]

[4] We use the WLE score since we do not compare between countries.

6 Empirical Evidence

In this section the empirical analysis is presented. We test the following hypotheses:

Hypothesis 1 An increase in the length of parents' total birth related leave is beneficial for children, i.e. has a measurable, positive impact on their long term cognitive outcomes. Furthermore, it has a positive impact on enrollment in secondary education.

Hypothesis 2 An increase in the length of parents' total birth related leave will not have a (long term) impact on children when the maternity leave is already long (more than 3 months).

We expect to find support for Hypothesis 1 but descriptive evidence suggests surprisingly a weak negative relationship between length of birth related leave and the child's cognitive outcome, see Figure 5. The mean reading score for June is higher than for July which is against our prior expectations. We also see seasonality (age) effects since there are hump shapes and a general decreasing trend. The evidence in Figure 5 is descriptive so it may be the case that the results change when we control for family and background characteristics. This will be done using the method of regression discontinuity design as mentioned earlier.

Finally, we need to make some robustness checks. First, we can use children born 15 days on either side of July 1st instead of a month, and secondly, we can use RD to estimate the effect of the policy reform in 1985. Finding an effect of the reform in 1985 makes it very unlikely that there will not be an effect of the reform in 1984 since the increase in parents' birth related leave by July 1st, 1985 is smaller than the increase in 1984 both in absolute values and by the percentage increase.

We will also estimate differences-in-differences using the sample of children born in June and July 1983 as the control group for both the 1984 and 1985-reform.

6.1 Estimation Results

[TBW - waiting for the data]

7 Conclusion

[TBW]

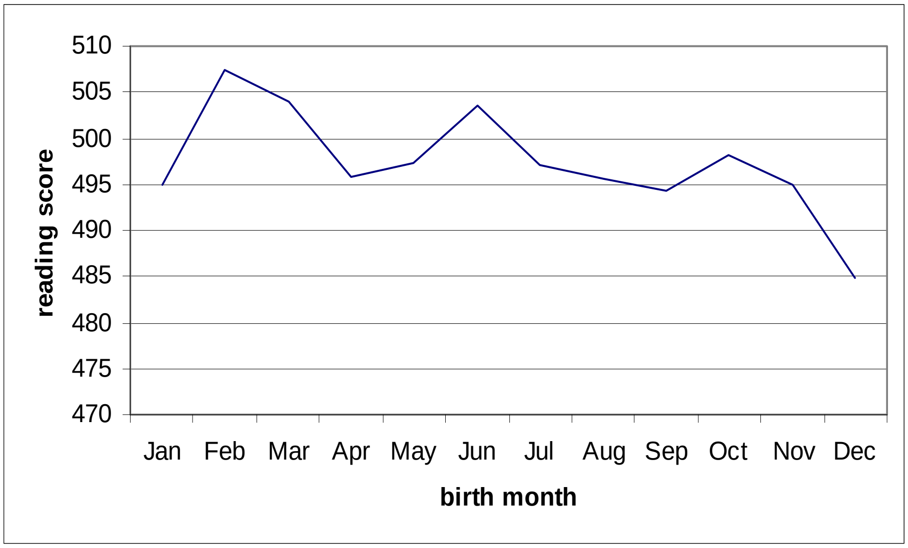

Figure 5: PISA-2000 reading scores for Danish children born in 1984.

References

Angrist, J. D. and V. Lavy (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics 114*(2), 533–575.

Baker, M. and K. Milligan (2005). How does job-protected maternity leave affect mothers' employment and infant health? *NBER Working Paper Series No. 11135*.

Baum, C. L. (2003). Does early maternal employment harm child development? An analysis of the potential benefits of leave taking. *Journal of Labor Economics 21*(2), 409–448.

Berger, L. M., J. Hill, and J. Waldfogel (2005). Maternity leave, early maternal employment and child health and development in the US. *The Economic Journal 115*, F29–F47.

Chatterji, P. and S. Markowitz (2005). Does the length of maternity leave affect maternal health? *Southern Economic Journal 72*(1), 16–41.

Gregg, P., E. Washbrook, C. Propper, and S. Burgess (2005). The effects of a mother's return to work decision on child development in the UK. *The Economic Journal 115*, F48–F80.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Lalive, R. and J. Zweimüller (2005). Does parental leave affect fertility and return-to-work? evidence from a "true natural experiment". *IZA Discussion Paper No. 1613*.

OECD (2002). PISA 2000. Technical report, OECD, Paris.

Pylkkänen, E. and N. Smith (2004). The impact of family-friendly policies in denmark and sweden on mothers' career interruptions due to childbirth. *IZA Discussion Paper No. 1050*.

Ruhm, C. J. (2000). Parental leave and child health. *Journal of Health Economics 19*(6), 931–960.

Ruhm, C. J. (2004). Parental employment and child cognitive development. *Journal of Human Resources 39*(1), 155–192.

Ruhm, C. J. (2005). Maternal employment and adolescent development. *IZA Discussion Paper No. 1673*.

Thistlethwaite, D. L. and D. T. Campbell (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology 51*, 309–317.