

Havnes, Tarjei; Mogstad, Magne

Working Paper

Is universal child care leveling the playing field? Evidence from non-linear difference-in-differences

IZA Discussion Papers, No. 4978

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Havnes, Tarjei; Mogstad, Magne (2010) : Is universal child care leveling the playing field? Evidence from non-linear difference-in-differences, IZA Discussion Papers, No. 4978, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/36832>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 4978

**Is Universal Child Care Leveling the Playing Field?
Evidence from Non-Linear Difference-in-Differences**

Tarjei Havnes
Magne Mogstad

May 2010

Is Universal Child Care Leveling the Playing Field? Evidence from Non-Linear Difference-in-Differences

Tarjei Havnes

ESOP, University of Oslo

Magne Mogstad

Statistics Norway,

ESOP and IZA

Discussion Paper No. 4978

May 2010

IZA

P.O. Box 7240

53072 Bonn

Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Is Universal Child Care Leveling the Playing Field? Evidence from Non-Linear Difference-in-Differences*

Advocates of a universal child care system offer a two-fold argument: Child care facilitates children's long-run development, and levels the playing field by benefiting in particular disadvantaged children. Therefore, a critical element in evaluating universal child care systems is to measure the impact on child development in a way that allows the effects to vary systematically over the outcome distribution. Using non-linear DD methods, we investigate how the introduction of large-scale, publicly subsidized child care in Norway affected the earnings distribution of exposed children as adults. We find that mean impacts miss a lot: While child care had a small and insignificant mean impact, effects were positive over the bulk of the earnings distribution, and sizable below the median. This is an important observation since previous empirical studies of universal child care have focused on mean impacts. We further demonstrate that the essential features of our empirical findings could not have been revealed using mean impact analysis on typically defined subgroups. This is because the intragroup variation in the child care effects is relatively large compared to the intergroup variation in mean impacts.

JEL Classification: J13, H40, I28, D31

Keywords: universal child care, child development, non-linear difference-in-differences, heterogeneity, distributional effects

Corresponding author:

Magne Mogstad
Research Department
Statistics Norway
P.O. Box 8131 Dep.
N-0033 Oslo
Norway
E-mail: magne.mogstad@ssb.no

* Thanks to Rolf Aaberge, Erling Barth, Nabanita Datta Gupta, Halvor Mehlum, Hilary Hoynes, Kalle Moene, Mari Rege and Kjetil Telle, as well as participants at a number of seminars and conferences for useful comments and suggestions. Financial support from the Norwegian Research Council (194347/S20) is gratefully acknowledged. The project is also part of the research activities at the ESOP center at the Department of Economics, University of Oslo. ESOP is supported by The Research Council of Norway.

1 Introduction

The increased demand for child care associated with the rise of maternal employment is attracting the attention of policy makers and researchers alike. Indeed, access to child care has gone up in many developed countries over the last years (OECD, 2004), and there is a heated debate about a move towards subsidized, universally accessible child care or pre-school, as offered in the Scandinavian countries. For example, the European Union's Presidency formulated in 2002 as a policy goal "*to provide childcare by 2010 to at least 90% of children between 3 years old and the mandatory school age and at least 33% of children under 3 years of age*" (EU, 2002, p. 13). In the US, the so-called '*Zero to Five Plan*' of US President Obama aims at making states move towards voluntary universal preschool. Advocates of such a universal child care system argue that it is important in facilitating children's long-run development. Moreover, it is claimed to be leveling the playing field by benefiting especially disadvantaged children (see e.g. Karoly et al. 2005). Therefore, a critical element in evaluating universal child care systems is to measure the impact on children's outcomes in a way that allows for heterogeneous treatment effects. That is the focus of this study.

In this paper, we provide first evidence on the distributional effects of universal child care on children's outcomes. As in Baker et al. (2008) and Havnes and Mogstad (2009), universal child care is taken to mean large-scale, publicly subsidized child care arrangements open for everyone; not that all children were in fact using child care. Specifically, we analyze how a Norwegian child care reform from late 1975 affected the earnings distribution of exposed children as adults. The reform assigned responsibility for child care to local governments and increased federal subsidies, which immediately generated large variation in child care coverage for children 3–6 years old, both across time and between municipalities.¹ As discussed below, formal child care both before the reform period and during the expansion was severely rationed, with informal care arrangements (such as friends, relatives, and unlicensed care givers) servicing the excess demand. In our analysis, we will focus on years immediately after the reform, when child care coverage increased from 10 to 28 percent, most likely reflecting an abrupt slackening of constraints on the supply side, rather than a spike in the local demand.

Our empirical analysis utilizes high-quality panel data from administrative registers covering the entire resident population and all licensed care givers in Norway. To identify the child care effects, we exploit that the supply shocks to formal care were larger in some areas than others. Specifically, we use standard (DD) methods to estimate mean impacts, comparing the adult earnings for 3 to 6 year olds before and after the reform, from municipalities where child care expanded a lot (i.e. the treatment group) and munic-

¹Throughout this paper, child care coverage rates refer to formal care, including publicly and privately provided child care institutions as well as licensed care givers, all eligible to subsidies from the government.

ipalities with little or no increase in child care coverage (i.e. the comparison group). In a similar vein, we use non-linear DD methods to map out the child care effects on the entire earnings distribution of exposed children as adults. This allows us to move beyond mean impacts, examining whether the effect of child care is constant across the distribution, or whether it leads to larger changes in certain parts of the distribution.

In our empirical analysis, we use two different non-linear DD methods. First, we take the method proposed by Firpo et al. (2009) for estimating unconditional quantile treatment effects (under the conditional independence assumption) to a DD framework, controlling also for unobserved time and group effects. Their method turns the difficult estimation problem of estimating the treatment effect on *unconditional* quantiles of the outcome distribution, into the simple estimation problem of estimating the treatment effect on the probability of being above a certain threshold of the outcome distribution. We further provide empirical results using an alternative approach to analyze the distributional effects of policy changes within a DD framework: the quantile DD model.² Both methods estimate the counterfactual distribution, which we use to map out the child care effects on the long-run earnings distribution of the children. It's heartening to find that even though the identifying assumptions are different, the results are quite similar. To further increase the confidence in our empirical strategies, we run a battery of specification checks.

The insights from our empirical results may be summarized with three conclusions. First, mean impacts miss a lot, concealing major heterogeneity: While child care had a small and insignificant mean impact, effects were positive over the bulk of the earnings distribution, and sizable below the median. This is an important observation since previous empirical studies of universal child care have focused on mean impacts.³ Second, the estimated heterogeneity in child care effects is consistent with predictions from economic theory. In particular, formal child care is predicted to reduce the importance of family background for child development by serving as a substitute for parental care or informal care arrangements, in end effect leveling the playing field. Third, the essential features of our empirical findings could not have been revealed using mean impact analysis on typically defined subgroups. This is because the intragroup variation in the child care effects is relatively large compared to the intergroup variation in mean impacts.⁴

In sum, our study shows that non-linear DD methods can play a useful role in assessing policy changes, when only non-experimental data is available and theory predicts heterogeneous treatment effects. Our analysis also serves as an example of how the pivotal argument in favor of universal child care might not be that it, on average, improves the

²See Athey and Imbens (2006) for a detailed discussion of the quantile DD method, and Poterba et al. (1995) and Meyer et al. (1995) for applications.

³For a recent review, see Almond and Currie (2010).

⁴This finding echoes the conclusion drawn in Bitler et al. (2006), who evaluate the labor supply responses to the Jobs First welfare reform.

long-run prospects of children, but rather that it levels the playing field.

This paper proceeds as follows. Section 2 discusses our study in relation to previous research on child care and child development. Section 3 describes the 1975 child care reform and the succeeding expansion in child care, before outlining a parsimonious theoretical framework making predictions about the effects of the policy changes on child development. Section 4 outlines the empirical strategy and Section 5 presents our data. Section 6 presents the main empirical findings, and reports results from specification checks. Section 7 concludes with a discussion of policy implications.

2 Child care and child development

Recent research from a number of fields suggests that investments in early childhood have high returns, especially for disadvantaged children (Knudsen et al., 2006). Studies in neuroscience and development psychology indicate that learning is easier in early childhood than later in life (Shonkoff and Phillips, 2000). In the economics literature, Becker (1964) points out that the returns to investments in early childhood are likely to be relatively high, simply because of the long time to reap the rewards. Taking this argument one step further, Carneiro and Heckman (2004) argue that investments in human capital have dynamic complementarities, implying that learning begets learning.

On this background, Currie (2001) suggests that governments should aim to equalize initial endowments through early childhood development, rather than compensate for differences in outcomes later in life. The role of governments in facilitating child development is particularly important, both from positions on equity and efficiency, if families underinvest in early childhood due to market failures such as liquidity constraints, information failures, and externalities (Gaviria, 2002).

Child care institutions are important arenas for child development, and expanding child care coverage is an explicit goal in many countries. This has been motivated by evidence showing that early childhood educational programs can generate learning gains in the short-run and, in many cases, improve the long-run prospects of children from poor families.⁵ While the results from these studies are encouraging, the programs evaluated were unusually intensive and involved small numbers of particularly disadvantaged children from a few cities in the US. A major concern is therefore that this evidence may tell us little about the effects of universal child care systems offered to the entire population (Baker et al., 2008). Nonetheless, it has fuelled an increasing interest in universal provision of child care as a means of advancing child development and improving children's long-run outcomes.

⁵The Perry Preschool and Abecedarian programs are commonly cited examples of how high-quality preschool services can improve the lives of disadvantaged children. See Barnett (1995) and Karoly et al. (2005) for surveys of the literature.

Our paper contributes to a small but rapidly growing literature on the effects of universal child care programs. Almost all the evidence is limited to short-run outcomes and the findings are mixed. Loeb et al. (2007), for instance, find that pre-primary education in the US is associated with improved reading and mathematics skills at primary school entry. However, Magnuson et al. (2007) suggest that these effects dissipate for most children by the end of first grade. Positive effects of child care on children's short-run outcomes are also found by Gormley and Gayer (2005), Fitzpatrick (2008), Melhuish et al. (2008), and Berlinski et al. (2008, 2009). On the other hand, Baker et al. (2008) analyze the introduction of subsidized, universally accessible child care in Quebec, finding no impact on children's short-run cognitive skills but substantial negative effects on children's short-run non-cognitive development. These negative effects echo the results in Herbst and Tekin (2008), while Datta Gupta and Simonsen (2007) find that compared to home care, being enrolled in preschool does not lead to significant difference in child non-cognitive outcomes.

While the evidence on short-run effects of universal child care programs is of interest, a crucial question is whether these effects persist, and perhaps are amplified, over time. As noted by Baker et al. (2008), negative short-run effects could reflect that children have difficulties in their first interactions with other children. In that case, child care attendance may expose children to these costs earlier on, so that they are better prepared for attending school. In addition, evidence from early intervention programs targeting particularly disadvantaged children suggests that even though the short-run gains in test-scores tended to dissipate over time, there were strong and persistent impacts on long-run outcomes (Heckman et al., 2006). Havnes and Mogstad (2009) and Cascio (2009) circumvent these issues by investigating the impact of universal child care on adult outcomes that are of intrinsic importance. In doing so, they also avoid reliance on test scores and changes in test scores that have no meaningful cardinal scale (see Cunha and Heckman, 2008).

Havnes and Mogstad (2009) find that universal child care in Norway had strong positive impacts on children's educational attainment and labor market participation as adults. Their subsample analysis indicates that girls and children with low educated parents benefit the most from child care. In terms of mechanisms, they find that the increase in formal child care largely displaced informal care, without much effect on mother's labor force participation. Cascio (2009) uses data from four decennial censuses to analyze the introduction of public preschools in the US. Using a cohort-based design, her baseline specification suggests that white children born after the reform in states that began funding kindergartens, largely in the South, were less likely to drop out of high-school. Yet she finds no effect on several other outcomes, like employment, college attendance, and earnings. Nor does she find any effects for blacks. She interprets the general lack of program effects as a result of (i) the low-intensity nature of the program, (ii) significant crowding out of participation in federally-funded programs, such as Head Start, and (iii)

cut-backs in state expenditure on schools to fund kindergartens.

However, the policy debate on universal child care policies is not restricted to whether child care, on average, improves child development: Distributional considerations also come to play. As emphasized by Almond and Currie (2010), formal child care is likely to reduce the importance of family background for child development by serving as a substitute for parental care or informal care arrangements, in end effect leveling the playing field. A concern is therefore that the estimated mean impacts may average together effects of different sign and magnitude, possibly obscuring the extent of child care's effects. The focus of our study is to address this concern. Specifically, we follow Havnes and Mogstad (2009) in considering a Norwegian child care reform from late 1975, introducing subsidized, universally accessible child care. We also use the same data, administrative registers covering the entire population. Our point of departure is to use non-linear DD methods to map out the child care effects on the entire earnings distribution of exposed children as adults, rather than using a standard DD approach to estimate mean impacts like previous studies.

3 Background and Theoretical Framework

In this section, we describe the Norwegian child care system before and after the 1975 reform,⁶ before outlining a parsimonious theoretical framework making predictions about the effects of the policy changes on child development.

The child care reform. In the post-WWII years in Norway, the gradual entry on the labor market of particularly married women with children, caused growing demand for out-of-home child care. In a survey from 1968, when child care coverage was less than five percent, about 35% of mothers with 3 to 6 year olds stated demand for formal child care (NOU, 1972). In the same survey, only 34% of the latter group of respondents stated that they were in fact using out-of-home child care on a regular basis. Out of these, just 14 percent were in formal child care, while more than 85 percent were using informal arrangements.⁷

The severe rationing of formal child care acted as a background for political progress towards public funding of child care.⁸ In the early 1950s, grants and subsidized loans were temporarily made available for construction and refurbishment of child care institutions, and their operation was regulated by law in 1954. Federal subsidies to formal child

⁶The description of the reform draws heavily on Havnes and Mogstad (2009). See also Leira (1992, ch. 4) for a detailed survey of the history of Norwegian child care policies since WWII.

⁷Relatives stand out as the largest group of informal care givers at 35 percent, followed by play parks at 20 percent, maids at 14 percent, other unlicensed care givers at 10 percent, and finally more irregular arrangements (such as neighbors and friends) at 7 percent (NOU, 1972).

⁸See Leira (1992, ch. 4) and The Norwegian Ministry of Children and Family Affairs (1998) for detailed surveys of the history of Norwegian child care policies since WWII.

care were assigned a permanent post on the national budget in 1962, and increased over the subsequent ten years from a modest USD 50 per child care place to a maximum of more than USD 1,200 annually.⁹ The child care subsidies were contingent on a federally determined maximum price to be paid by the parents, which in 1972 was about USD 215 per month for full time care (NOU, 1972).

In 1972, the Norwegian government presented the Kindergarten White Paper (NOU, 1972), proposing radical changes in public child care policies. To (i) create positive arenas for child development, (ii) free labor market reserves among mothers, and (iii) lessen the burden on parents and relieve stress in the home, it was argued that child care should be made universally available. This marked a strong shift in child care policies, from focusing on children with special needs (in particular disabled children and children from disadvantaged families) to a focus on a child care system open to everyone.

In June 1975, The Kindergarten Act was passed by the Norwegian parliament with broad bipartisan political support. It assigned the responsibility for child care to local municipalities, but included federal provisions on educational content, group size, staff skill composition, and physical environment. By increasing the level of federal subsidies for both running costs in general and investment costs for newly established institutions, the government aimed at quadrupling the number of child care places to reach a total of 100,000 by 1981.¹⁰

In the years following the reform, the child care expansion was progressively rolled out at a strong pace, with federal funding more than doubling from USD 34.9 million in 1975 to 85.8 million in 1976, before reaching 107.3 million in 1977.¹¹ This implied an increase in the federal coverage of running costs from about 10% in 1973 to 17,6% in 1976, and further to 30% in 1977. From 1976, newly established child care places received additional federal funds for a period of five years. Municipalities with relatively low child care coverage rates were awarded 60% more subsidies, whereas other municipalities were awarded 40% extra.

Altogether, the reform constituted a substantial positive shock to the supply of formal child care, which had been severely constrained by limited public funds. In succeeding years, the previously slow expansion in subsidized child care accelerated rapidly. From a total coverage rate of less than 10% for 3 to 6 year olds in 1975, coverage had shot up above 28% by 1979. Over the period, a total of almost 38,000 child care places were established, more than a doubling from the 1975-level. By contrast, there was almost no child care coverage for 1 and 2 year olds during this period. Figure 1 draws child care

⁹Throughout this paper, all monetary figures are fixed at 2006-level. For the figures expressed in US dollars, we have used the following exchange rate: NOK/USD = 6.5.

¹⁰In addition, the price-setting was delegated to local municipalities, abolishing the federally determined maximum parental price for child care subsidies. However, Gulbrandsen et al. (1981) report survey data suggesting that the maximum price to be paid by the parents actually changed little in the years following the reform, and formal child care remained rationed well into the 90s

¹¹Source: National budgets 1975/76 through 1978/79.

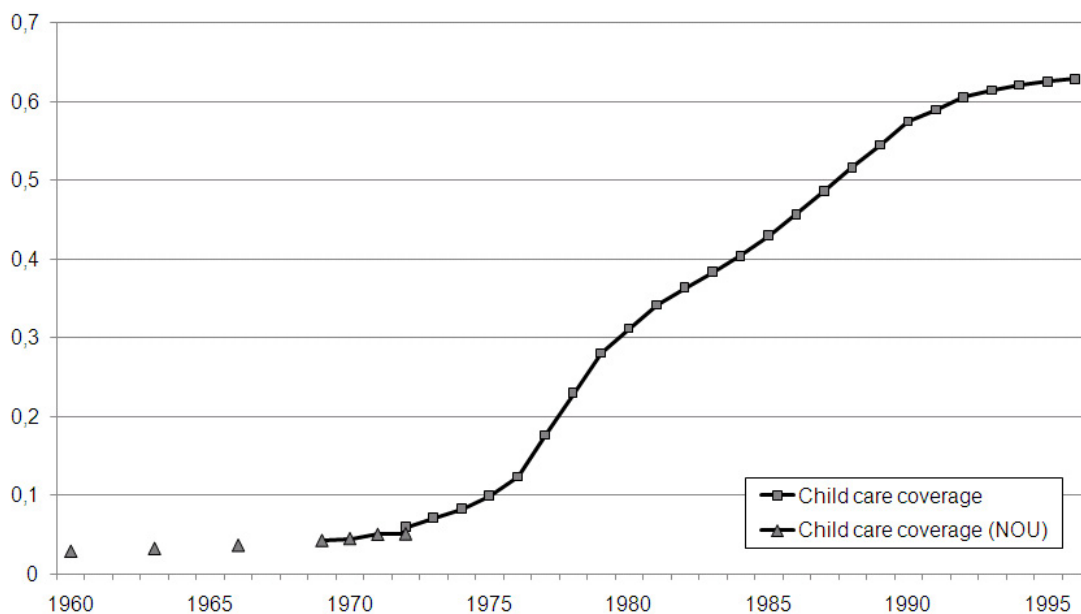


Figure 1: Child care coverage rate in Norway 1960–1996 for children 3–6 years old. Sources: Administrative data for 1972–1996. Data for 1960–1972 from NOU (1972), Table II.1.

coverage rates in Norway from 1960 to 1996 for 3 to 6 year olds. As is apparent from the figure, there has been strong growth in child care coverage rates since 1975, particularly in the early years. In our analysis, we will focus on the early expansion, which likely reflects an abrupt slackening of constraints on the supply side, rather than a spike in the local demand.

We might worry about confounding the estimated child care effects with other reforms or changes taking place in the same period. However, we have found no significant reforms or breaks in trends that could be of concern for our estimations. An extension in maternity leave implemented in 1977 did not affect the children in our sample directly, but could potentially influence family size, which could in turn matter for child development. However, the reform was nationwide, and should be controlled for by cohort fixed-effects. In addition, our rich set of controls may pick up potentially remaining effects of this policy change. Importantly, there were no significant changes in the Norwegian educational policies affecting the cohorts of children we consider. On the contrary, Norway was known for its unified public school system based on a common national curriculum, rooted in a principle of equal rights to high-quality education, regardless of social and economic background or residency. This is mirrored in very similar expenditure levels per student across municipalities and virtually no private schools.¹²

The Organization of Formal Child Care. To interpret our results, we must understand the type of care we are studying. The Ministry of Consumer Affairs and Administration was responsible for overall regulation of formal child care. Specifically, the Kindergarten

¹²See Telhaug et al. (2006), Volckmar (2008), and Havnes and Mogstad (2009) for an in-depth discussion of the Norwegian educational system relevant for the cohorts of children we consider.

Act regulated the authorisation, operation and supervision of formal child care institutions. The act defined formal child care institutions as care and educationally oriented enterprises for pre-school children, where an educated preschool teacher was responsible for the education. Formal child care institutions were run either by the municipalities or by firms, public institutions or private organisations, under the approval and monitoring of local authorities in the municipality. Table 1 reports child care institutions by owner biannually from 1975 through 1981, and shows the strong growth in municipal and cooperative child care centers. Over the period, the share of private centers decreased from 28 to 22 percent, driven almost entirely by a decline in the share of centers run by private organizations.

Regardless of ownership, formal child care institutions were required to satisfy federal provisions on educational content and activities, group size, staff skill composition and physical environment. The Kindergarten Act specified regulations, and guidelines were formulated for activities and content. To be eligible for subsidies, institutions were obliged to meet the requirements and follow the guidelines. To secure opportunities for parental involvement and promote cooperation between staff and parents, the Kindergarten Act required that every institution must have a parent council and a coordinating committee. Local authorities were required by law to monitor the fulfillment of these federal provisions.

As discussed above, formal child care institutions were financed jointly by the federal government, municipalities and parents. All approved institutions received subsidies for running and establishment costs from the federal government. Subsidies were determined on the basis of the number and age of children, and the amount of time they spent in formal child care. In general, formal child care institutions were open during normal working hours. All children were eligible, and open slots were in general allocated according to length of time on the waiting list and age. Only under special circumstances could a child gain priority on the waiting list.

Table 1: Child care institutions by ownership structure

	1975	1977	1979	1981
Private (%)	28.4	26.7	26.3	21.9
Municipality (%)	48.6	45.4	46.9	51.2
Church (%)	7.3	8.0	8.6	8.6
Cooperatives (%)	5.6	8.2	9.7	10.0
No. of child care institutions	880	1,469	2,294	2,754
No. of children in child care (3–6 y.o.)	25,536	43,239	63,218	73,152
Coverage rate (3–6 y.o., %)	10.0	17.6	28.1	34.2

Notes: Private ownership indicates ownership by a private firm, organisation or foundation. Cooperatives are parental or residential. Categories not reported are ownership by state, regions and other.

Every formal child care institution had to be run by an educated pre-school teacher responsible for day-to-day management. The pre-school teacher education is a college degree, including supervised practice in a formal child care institution. Through his or her position and training, this head teacher was responsible for ensuring satisfactory planning, observation, collaboration and evaluation of the work. The head teacher was also in charge of staff guidance, as well as collaboration with parents and local authorities, such as health stations, child welfare services and educational/psychological services. In addition, formal child care institutions were required to have at least one educated pre-school teacher per 16 children aged 3–6. Teachers typically worked closely with one or two assistants, and were responsible for the educational programmes in separate groups and for day-to-day interaction with parents. There were no educational requirements for assistants.

In terms of educational content, a social pedagogy tradition dominated the child care practices, according to which children were supposed to develop social, language and physical skills mainly through play and informal learning.¹³ The informal learning was typically carried out in the context of day-to-day social interaction between children and staff, in addition to specific activities for different age groups.

Overall, formal child care in Norway (along with other Nordic countries) was characterized by relatively high expenditure levels per child compared to large-scale programs in other countries. For example, the average yearly expenditure for a slot in formal child care was approximately USD 6,600.¹⁴ This is, for instance, substantially higher than the expenditures for the Head Start Program in the US aimed at low-income families, which cost around USD 5,000 per year (Currie, 2001). The high expenditure levels were mirrored in fairly extensive requirements to qualifications of child care staff and physical environment, as well as a relatively low number of children per staff. For example, the average staff–child ratio was about 1:8 in 1977. In comparison, in the US and Canada, the corresponding ratio is 1:12, in Spain 1:13, and France 1:19 (see Datta Gupta and Simonsen, 2007).

Theoretical framework. As discussed in Currie and Almond (2010), formal child care is likely to reduce the importance of family background for child development by serving as a substitute for parental care or informal care arrangements, in end effect leveling the playing field. Specifically, our child care reform may be interpreted as a subsidy to parents for choosing out-of-home care of a particular quality, generating at least one and

¹³The social pedagogy tradition to early education has been especially influential in the Nordic countries and Central-Europe. In contrast, a so-called pre-primary pedagogic approach to early education has dominated many English and French-speaking countries, favoring formal learning processes to meet explicit standards for what children should know and be able to do before they start school.

¹⁴Estimated annual budgetary cost per child care place from NOU (1972) is about USD 5,400 per child 3–6 years old. In addition, investment costs are estimated at about USD 12,000 per child care place, adding USD 1,200 to the annual cost if written down over ten years.

possibly two convex kinks in the family's budget constraint. Figure 2 illustrates this point by drawing family budget frontiers between child quality and parental consumption before and after the reform. The parents trade off their own consumption, c , and child quality, q , given their budget constraint. For instance, parents could invest in children by decreasing labor supply, by paying for higher quality out-of-home care, or by purchasing child goods. The budget frontier of *feasible* combinations of consumption and child quality then resembles a standard production possibilities frontier.¹⁵

In Panel (a) of Figure 2, subsidized child care is not available. Trading off child quality with their own consumption, parents optimize by choosing a point in the tangency of an indifference curve and the budget frontier. Panel (b) and (c) draw two examples of alternative budget frontiers (dotted curves), in a situation with subsidized child care. The decision to (apply for and) take up a child care place is discrete. In reality, parents could typically choose to pay for either a full or a half day of care. For simplicity, we assume full discreteness where parents either use formal child care or they don't. Because parents could choose not to spend money on formal child care and get higher consumption, the frontiers using formal child care will be below the frontier without formal child care at the c -axis. Further, if parents choose to purchase formal child care, they must be better off. That is, for parent's to accept the offer of subsidized formal care, the budget frontiers with formal child care must be outside the frontier without formal child care, generating at least one convex kink where the latter curve is steeper than the former curve at the point of intersection. Introduction of formal child care must therefore have a positive effect on child quality around this point. This is illustrated in Panels (b) and (c): Parents who were previously located in a point like A will now optimally locate in a point like A' , where child quality is higher. Their children will benefit from the reform, at a decreasing rate in their pre-reform q .

If the maximum feasible child quality is lower with formal child care than without, the two frontiers will intersect again, generating a second convex kink.¹⁶ However, in this case the frontier with formal child care will be steeper than the curve without formal child care. The introduction of formal care will, therefore, have an unambiguously negative effect on child quality around this point. This is illustrated in Panel (b): Parents who were previously located in a point like B will now optimally move to point B' , trading off child quality for parental consumption. The reform will be detrimental for their children's quality, and at an increasing rate in their pre-reform q . In comparison, Panel (c) considers

¹⁵Formally, with fixed wages w and unit consumer prices and time endowment, the budget constraint is $w(1 - h) \leq c + p(q)$, where h is home care, and $p(q)$ reflects the cost-minimizing price of providing child quality q . With child quality produced from a standard child production function, $p(q)$ should be increasing and convex in quality.

¹⁶The existence of a second intersection depends on the possibility and efficiency of topping up formal care investments with market goods or informal care. While using formal care should exclude additional investments during a substantial portion of the day, suggesting lower maximal quality, parents can be thought to be somewhat richer since formal care is subsidized, pulling in the opposite direction.

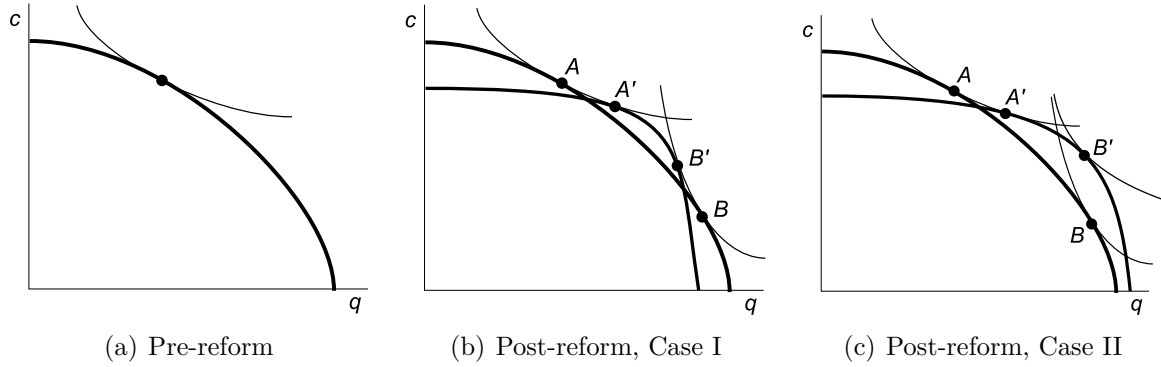


Figure 2: Family budget frontiers between child quality and parental consumption before and after the reform.

Notes: Panel (a) displays budget frontier without subsidized formal child care. Panels (a) and (b) display budget frontiers with and without subsidized formal child care. In the panel (b), the maximum feasible child quality is lower with formal child care than without. In the panel (a), the maximum feasible child quality is lower without formal child care than with.

a situation where the two budget frontiers do not intersect a second time. In this case, the overall effect is ambiguous, and depends on parental preferences. Specifically, the reform effect on children from families who were previously investing heavily in child quality, located in points like B , will be ambiguous and depend on income and substitution effects. Finally, parents that prior to the reform choose a q far left of A or far right of B will not be affected by the child care reform, because they will not use any formal care offered.

In sum, the predicted effects on children's outcome distribution of the child care reform are heterogeneous: There should be no change at the bottom, increases in the lower and middle parts, decreases at the upper part, and perhaps no change at the very top. In such a situation, the mean impact would average together effects of different sign and magnitude across the distribution, possibly obscuring the extent of child care's effects.

4 Empirical strategy

To estimate the effects of universal child care on children's earnings as adults, we follow the previous literature closely in using a DD framework. In particular, our identification exploits that the supply shocks to formal child care were larger in some areas than others. Below, we describe our treatment definition, before specifying our standard and non-linear DD approaches. As always in policy evaluation using non-experimental data we cannot completely guard against such omitted variables bias. Yet to increase the confidence in our identification strategies, we run a battery of specification checks which are discussed after the main results, in Section 6.

Treatment definition. Our main empirical strategy is the following: We compare adult earnings for children who were 3 to 6 year olds before and after the reform, from municipalities where child care expanded a lot (i.e. the treatment group) and municipalities

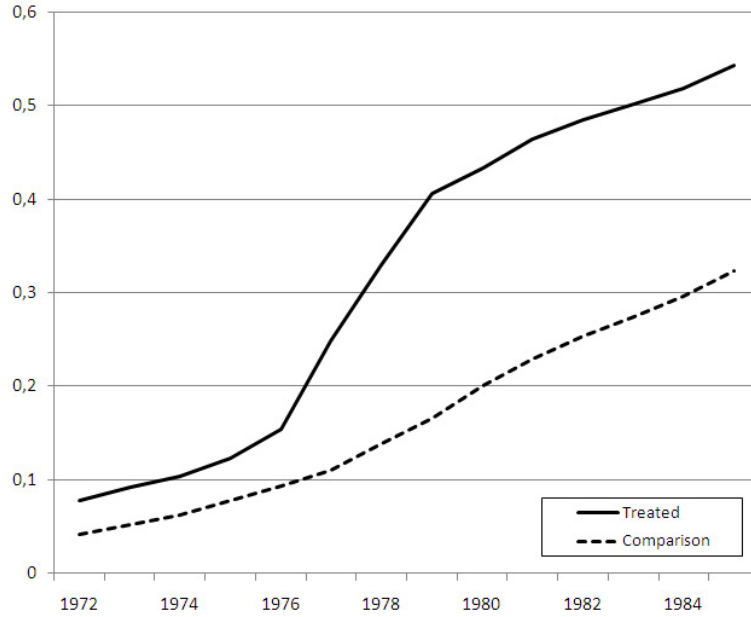


Figure 3: Child care coverage rates 1972–1985 for 3–6 year olds in treatment and comparison municipalities

Notes: Treatment (comparison) municipalities are above (below) the median in child care coverage growth from 1976 to 1979.

with little or no increase in child care coverage (i.e. the comparison group).

The child care expansion started in 1976, affecting the *post-reform cohorts* born 1973–1976 with full force, and to a lesser extent the *phase-in cohorts* born 1970–1972. The *pre-reform cohorts* consist of children born in the period 1967–1969. We consider the period 1976–1979 as the child care *expansion period*. In the robustness analysis, we take several steps to ensure that our results are robust to the exact child care coverage cut-off, defining treatment and comparison groups.

To define the treatment and comparison group, we order the municipalities according to the percentage point increase in child care coverage rates from 1976 to 1979. We then separate the sample at the median, letting the upper half constitute the *treatment municipalities* and the lower half the *comparison municipalities*. Figure 3 displays child care coverage before and after the 1975 reform in treatment and comparison municipalities (weighted by population size). The graphs move almost in parallel before the reform, while the child care coverage of the treatment municipalities kinks heavily after the reform. This illustrates that our study compares municipalities that differ distinctly in terms of changes in child care coverage within a narrow time frame. In the robustness analysis, we consider whether variations in treatment intensity affects our results, by changing the child care coverage cut-off defining the treatment and comparison municipalities.

4.1 Mean impact estimation

When estimating mean impacts in the population of children as a whole and in subgroups by child and parental characteristics, we use the same DD specification as Havnes and Mogstad (2009). Our baseline regression model, estimated by OLS over the sample of children born during the period 1967–1976, can be defined as

$$Y_{ijt} = \psi_t + \gamma_1 \text{Treat}_i + \gamma_2 (\text{Treat}_i \times \text{Phasein}_t) + \theta (\text{Treat}_i \times \text{Post}_t) + X'_{ijt} \beta + \epsilon_{ijt}, \quad (1)$$

where Y is the earnings in 2006, i indexes child, j indexes family, and t indexes the year the child turns 3 years old. The vector of covariates X includes parental age, their education when the child is 2 years old, their age at first birth, the number of older siblings (also capturing birth order) and relocation between municipalities within treatment/comparison area, the child’s sex and immigrant status, as well as municipality-specific fixed effects. The error term ϵ_{ijt} is clustered on the mother, allowing for dependence in the residuals of siblings.¹⁷ The dummy variable Treated_i is equal to 1 if child i lives in the treatment area, Phasein_t and Post_t are dummy variables equal to 1 when $t \in [1973, 1975]$ and $t \in [1976, 1979]$ respectively, while ψ_t are cohort-specific fixed effects.

The parameter of interest, θ , captures the average effect on children who reside in the treatment area in the post-reform period, of additional child care slots following the reform in the treatment municipalities compared to the comparison municipalities. There are two types of averaging underlying this average causal effect. First, there is averaging over the impacts on children from different municipalities in the treatment area. And second, there is averaging across the marginal effects of the additional child care slots.

Like in Baker et al. (2008) and Havnes and Mogstad (2009), we will interpret θ as the mean intention-to-treat effect (ITT), since our regression model estimates the reduced form impacts on all children from post-reform cohorts who reside in the treatment area. An advantage of the ITT parameter is that it captures the full reform impact of changes in both formal and informal care arrangements, as well as any peer effects on children who were not attending child care. However, since this parameter averages the reform effects over all children in the municipality, it reflects poorly the *size* of the child care expansion. To arrive at the treatment-on-the-treated (TT) effect, we follow Baker et al. and Havnes and Mogstad and scale the ITT parameter with the probability of treatment. Specifically, we divide the ITT parameter with the increase in child care coverage following the reform in the treatment group relative to the comparison group. For example, $TT = ITT/0.1785$ in our baseline specification. The TT parameter gives us the effect of child care exposure – per child care place – on children born in post-reform cohorts who reside in the treatment

¹⁷Bertrand et al. (2004) show that the standard errors in DD regressions may suffer from serial correlation. As their analysis demonstrates, we reduce the problem of serial correlation considerably by collapsing the time-series dimension into three periods: pre-reform, phase-in, and post-reform.

area. For ease of interpretation, we will throughout the paper report the TT parameters (and their standard errors).

4.2 Non-linear difference-in-differences estimation

To map out the child care effects on the earnings distribution of exposed children as adults, we use non-linear DD methods. Our main empirical strategy extends the method proposed by Firpo et al. (2009) for estimating unconditional quantile treatment effects (under the conditional independence assumption) to a DD framework, controlling also for unobserved time and group effects. To define this threshold DD estimator, let F_{gt} denote the earnings distribution in group g at time t , where $g = 1$ indicates the treatment group (and $g = 0$ the comparison group) and $t = 1$ indicates the post-reform cohorts (and $t = 0$ pre-reform cohorts). The estimator (without covariates) for the counterfactual post-reform distribution in the treatment group in the absence of the child care reform, \tilde{F}_{11}^T , can be defined as

$$\tilde{F}_{11}^T(y) = F_{10}(y) + [F_{01}(y) - F_{00}(y)].$$

The reform effect at a particular earnings level y , denoted $\tau^T(y)$, is then given by the difference between the actual (observed) earnings distribution and the estimated counterfactual distribution at this point, that is $\tau^T(y) = F_{11}(y) - \tilde{F}_{11}^T(y)$.

Following Firpo et al. (2009), the analogous regression model with covariates can be specified as

$$\Pr(Y_{ijt} \geq y) = \psi_i^y + \gamma_1^y Treat_i + \gamma_2^y (Treat_i \times Phasein_t) + \theta^y (Treat_i \times Post_t) \quad (2) \\ + X'_{ijt} \beta^y + \epsilon_{ijt}^y,$$

where X is the same set of controls as in the mean impact estimation. As shown by Firpo et al. (2009), this turns the difficult estimation problem of estimating the treatment effect on unconditional quantiles of the outcome distribution, into the easy estimation problem of estimating the treatment effect on the probability of being above a certain threshold of the outcome distribution. The reason is that, unlike the conditional quantile, the conditional mean averages up to the unconditional mean, due to the law of iterated expectations. As a result, the OLS estimate of θ in equation (2) gives the unconditional reform impact on the probability of earning at least y .¹⁸ In our empirical analysis, we estimate the reform effects at the earnings levels of all percentiles 1-99 in the pre-reform distribution in the treatment group.

¹⁸Our linear probability model will be the best least-squares approximation of the true conditional expectation function. As noted by Angrist (2001), if there are no covariates or they are discrete, as in our case, linear models are no less appropriate for limited dependent variables than for other types of dependent variables.

To examine the robustness of our results, we also report estimation results from another non-linear DD approach, the quantile DD method which is described in detail in Athey and Imbens (2006). As before, we estimate the counterfactual earnings distribution by taking the observed distribution among pre-reform cohorts from the treatment group and adjusting for the growth from pre-reform to post-reform cohorts in the earnings distribution of the comparison group. However, we adjust by using the change in the quantiles, rather than the change in population shares as in the threshold DD estimator. Specifically, the quantile DD estimator (without covariates) for the counterfactual post-reform distribution in the treatment group in the absence of the reform, \tilde{F}_{11}^Q , can be defined as

$$\tilde{F}_{11}^{Q^{-1}}(q) = F_{10}^{-1}(q) + [F_{01}^{-1}(q) - F_{00}^{-1}(q)]$$

The estimated reform effect at a particular quantile is then given analogously to the above, by $\tau^Q(q) = F_{11}^{-1}(q) - \tilde{F}_{11}^{Q^{-1}}(q)$.

The two non-linear DD methods are similar, but rely on different dimensions of the distributions in the identifying assumptions. To compare them more explicitly, Figure 4 draws the observed earnings distributions in 2006 separately for pre- and post-reform cohorts from treatment and comparison municipalities. Since post-reform cohorts are younger, it is not surprising that they have lower earnings than those from pre-reform cohorts. Further, it is evident that the pre-reform cohorts in the treatment and comparison groups have very similar earnings, until about the 70th percentile. At higher percentiles, earnings levels in the treatment group are somewhat higher than in the comparison group. This suggests that a direct comparison of children from treatment and comparison municipalities would be likely to overestimate the effect of the reform on these higher quantiles. The DD methods, however, achieve identification from the changes in these differences, so we are not concerned with differences between the treatment and comparison group in the earnings levels as such.

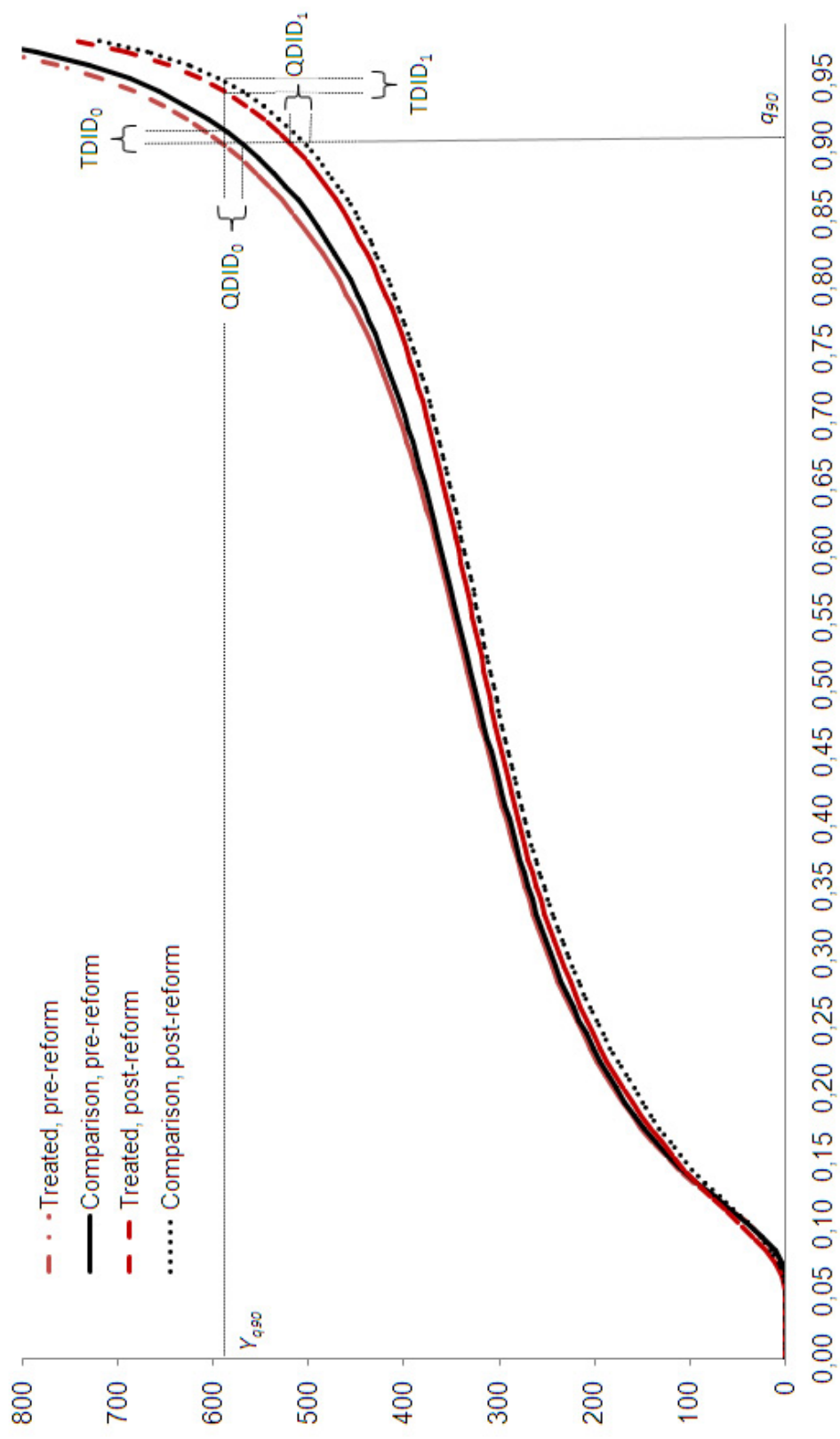


Figure 4: Earnings distribution in pre-reform and post-reform cohorts in treatment and comparison municipalities (NOK 1,000).

Notes: Pre-reform cohorts are born 1967–1969, and post-reform cohorts are born 1973–1976. Treatment (comparison) municipalities are above (below) the median in child care coverage growth from 1976 to 1979.

With threshold DD, we compare the horizontal differences in Figure 4. Consider, for example, the earnings level at the pre-reform 90th percentile, Y_{q90} . The threshold DD estimator compares the difference in population shares below this *level* of earnings among individuals from pre-reform cohorts, $TDID_0$, with the corresponding difference among children from post-reform cohorts, $TDID_1$. The reform effect at Y_{q90} is estimated as $TDID_1 - TDID_0$. Holding the *level* of earnings fixed, our threshold DD compares changes in the distributions around Y_{q90} , that is, we compare the population shares with earnings below a certain threshold. The identifying assumption boils down to the earnings growth from pre-reform to post-reform cohorts around this particular earnings level being the same in the treatment group as in the comparison group, in the absence of the reform.

In contrast, the quantile DD holds the *quantile* fixed, and compares changes in the level of earnings around this quantile. Considering again the 90th percentile marked in the figure, the quantile approach considers the difference in earnings at this quantile among pre-reform cohorts, $QDID_0$, and compares this to the difference in earnings among post-reform cohorts *at the same quantile*, $QDID_1$. The identifying assumption is that the growth in earnings from pre-reform to post-reform cohorts at this particular quantile would have been the same in the treatment group as in the comparison group, in the absence of the reform. However, as is evident at the 90th percentile in Figure 4, the earnings levels at these quantiles may be quite different between the treatment and comparison group. If the earnings levels at a given quantile differ between the treatment and comparison group and the growth in earnings depends on the earnings levels, it is likely that the identifying assumption of the quantile DD model is violated. In our view, this makes the quantile DD estimator somewhat less attractive than the threshold DD estimator.

Another advantage of the threshold DD estimator is that it is straightforward to control for differences between the treatment and comparison group in child and parental characteristics. Unlike conditional means, conditional quantiles do not average up to their unconditional counterparts. As a result, a conditional quantile regression cannot in general be used to identify the unconditional quantile treatment effects (see Firpo et al., 2009). To handle covariates in our quantile DD method, we therefore follow Firpo (2007) in using an inverse propensity score re-weighting procedure, with weights determined by the propensity for treatment depending on X . Appendix B describes our implementation of this procedure, and compares the results from the two non-linear DD methods, with and without covariates.

It should finally be noted that the non-linear DD methods identify the child care effect on the earnings distribution, which will, in general, differ from the distribution of child care effects, unless the effects are rank-preserving (see Heckman et al. 1997). However, our focus is on the distributional effects of the child care reform: For this purposes, the effect on the outcome distribution is what we are after. Also, any differences in the TT

parameters between subgroups and at different parts of the outcome distribution, may reflect both differences in child care take-up and potentially heterogeneous impacts of uptake, in line with the predictions from our theoretical framework. Adjusting for differences in take-up across the earnings distribution would require strong assumptions about how the child care reform shifts the ranks of the children. Making such assumptions is, however, unnecessary since our focus is on the distributional effects of the child care reform. In any case, we do not have data on child care use by child and parental characteristics.

5 Data

Like Havnes and Mogstad (2009), our data are based on administrative registers from Statistics Norway covering the entire resident population of Norway from 1967–2006. The data contains unique individual identifiers that allow us to match parents to their children. As we observe each child’s date of birth, we are able to construct birth cohort indicators for every child in each family. The family and demographic files are merged through the unique child identifier with data on his or her annual earnings in 2006. Our earnings measure includes all market income, from wages and self-employment. In a placebo test, we will also use data on adult height. Information on adult height is collected from military records and is only available for males, since military service is compulsory for men only. Before entering military service, medical and psychological suitability is assessed, including a measurement of height.¹⁹

Importantly, we also have administrative register data on all formal child care institutions and their location from 1972, reported directly from the institutions themselves to Statistics Norway. All licensed care givers are required to report annually the number of children in child care by age. Merging this data with the demographic files containing information about the total number of children according to age and residency, we construct a time series of annual child care coverage (by age of child) in each of the 418 municipalities. The reliability of Norwegian register data is considered to be very good, as is documented by the fact that they received the highest rating in a data quality assessment conducted by Atkinson et al. (1995).

We start with the entire population of children born during the period 1967–1976, who were alive and resident in Norway in 2006. This sample consists of 575,300 children, spanning these 10 birth cohorts. The choice of cohorts serves three purposes. Since our outcomes are measured in 2006 and 2007, the treated children are 30–33 and 31–34 years old at the time of measurement, which should be suitable when assessing adult earnings

¹⁹Eide et al. (2005) examine patterns of missing data in military records for males from the 1967–1987 cohorts. Of those, 1.2% died before 1 year and 0.9% died between 1 year of age and registering with the military at about age 18. About 1% of the sample of eligible men had emigrated before age 18, and 1.4% of the men were exempted because they were permanently disabled. An additional 6.2% are missing for a variety of reasons including foreign citizenship and missing observations.

prospects (see Haider and Solon, 2006; Böhlmark and Lindquist, 2006). Second, since treatment and comparison groups are defined by the expansion in child care from 1976 to 1979, the regional and time variation between the two groups breaks down as we move away from 1979. Indeed, the coverage rates do converge slowly after 1979. Finally, to ensure comparability of children before and after the reform, we do not want the cohorts to be too far apart.

We restrict the sample to children whose mothers were married at the end of 1975, which makes up about 92 percent of the above sample. The reason for this choice is that our family data does not allow us to distinguish between cohabitants and single parents in these years. As parents' family formation may be endogenous to the reform, we only condition on pre-reform marital status. To avoid migration induced by the child care reform, we also exclude children from families that move between treatment and comparison municipalities during the expansion period, which makes up less than 5 percent of the above sample. To minimize the sensitivity of the mean estimations to outliers, we exclude a small number of children with earnings above NOK 5 million. Finally, we drop a handful of children whose mother had a birth before she was aged 16 or after she was 49. Rather than excluding observations where information on parents' education is missing, we include a separate category for missing values. The education of the parents is measured when the child is 2 years old. The number of older siblings relates to children born to each mother. The final sample used in the estimations consists of 498,956 children from 318,345 families, which makes up about 87 percent of the children from each cohort. When interpreting our results, it is necessary to have these sample selection criteria in mind: We focus on children of married mothers, and our results do not speak to the literature on early childhood educational programs targeting special groups like children of single mothers – but these are not the central focus of the current policy debate.

Descriptive statistics. Table 2 shows mean earnings in 2006, as well as earnings levels at different percentiles. As is evident from the table, there are systematic changes in the differences in the earnings of the treatment and comparison group between the pre-reform, phase-in and post-reform cohorts. In our DD framework, this pattern is suggestive of significant effects of child care on children's earnings. In particular, Table 2 indicates positive effects at the lower and central parts and negative effects at the upper parts of the earnings distribution.

Our DD framework identifies the effects of child care by comparing the change in earnings before and after the reform of children residing in treatment and comparison areas. Substantial changes over time in the *differences* in the observable characteristics of the two groups may suggest unobserved compositional changes, calling our empirical strategy into question. Table 3 shows means of our control variables for characteristics of the child and the parents. It turns out that the treatment and comparison groups have fairly

Table 2: Descriptive statistics: Outcome variable

	– Level –	– Differences –		
	Treated Pre-reform	Treated Pre-reform	Comparison Phase-in	Comparison Post-reform
5th percentile	0	0	0	0
10th percentile	31,685	-13	3,211	8,081
25th percentile	215,559	2,735	3,424	9,352
50th percentile	328,825	3,601	4,083	6,346
75th percentile	431,591	7,650	8,713	7,668
90th percentile	588,319	20,891	18,489	14,401
95th percentile	718,938	30,293	23,727	19,812
Mean (SD)	343,361 (270,402)			

Notes: Pre-reform cohorts are born 1967–1969, phase in-cohorts are born 1970–1972, and post-reform cohorts are born 1973–1976. Treatment (comparison) municipalities are above (below) the median in child care coverage growth from 1976 to 1979. Outcome variable is defined in Section 5. Standard deviation in paranthesis.

similar characteristics. More importantly, there appears to be small changes over time in the relative characteristics of the two groups. We have also run our baseline specification replacing the dependent variable with each control variable. With the exception of relocation between municipalities within the treatment/comparison area, the results show small – and mostly insignificant – differences over time in the characteristics of children residing in the treatment and comparison areas. We have also performed all estimations excluding all families that relocate between municipalities within the treatment/comparison area. These estimations yield very similar results.

A concern in applying linear regressions is lack of overlap in the covariate distribution. As emphasized by Imbens and Wooldridge (2009), this can be assessed by the (scale-invariant) normalized difference measure. For each covariate, the normalized difference is defined as the difference in averages by treatment status, scaled by the square root of the sum of variances. Imbens and Wooldridge suggest as a rule of thumb that linear regression methods tend to be sensitive to the functional form assumption if the normalized difference exceeds one quarter. Table 3 displays normalized differences for our controls in curly brackets, indicating that lack of overlap should be of little concern for the estimated effects. In any case, we use a flexible specification for the vector of covariates, as discussed above.

Because we control for municipality-specific fixed effects, it is not necessary that the child care expansion is unrelated to municipality characteristics. However, if determinants of the expansion are systematically related to underlying trends in children’s potential outcomes, we may be worried about differences in the characteristics of treatment and comparison municipalities. For example, if expansive municipalities are aiming at counteracting a particularly negative trend in child development, or if they are taking some

Table 3: Descriptive statistics: Control variables

	– Level –	– Differences –		
	Treated Pre-reform	Pre-reform	Treated – Comparison Phase-in	Post-reform
Male	0.5069 [0.5000]	-0.0014 {-0.0020}	0.0036 {0.0051}	0.0017 {0.0024}
No. of older siblings	2.1319 [1.2343]	-0.0818 {-0.0456}	-0.0736 {-0.0432}	-0.1118 {-0.0718}
Mother’s age at first birth	23.3286 [4.0432]	0.5671 {0.1021}	0.5916 {0.1119}	0.6472 {0.1223}
Father’s age at first birth	26.5592 [5.2946]	0.4936 {0.0675}	0.4867 {0.0705}	0.5444 {0.0823}
Mother’s education when child 2 y.o.	9.6618 [2.0739]	0.2805 {0.0987}	0.2817 {0.0992}	0.3072 {0.1066}
Father’s education when child 2 y.o.	10.3715 [2.8162]	0.3730 {0.0971}	0.3787 {0.0995}	0.4044 {0.1065}
Immigrant	0.0566 [0.2311]	0.0110 {0.0355}	0.0165 {0.0535}	0.0162 {0.0534}
Relocated	0.0358 [0.1858]	-0.0016 {-0.0061}	0.0021 {0.0061}	0.0070 {0.0172}
No. of children (level)		77,933 – 87,832	74,182 – 83,621	84,052 – 91,406

Notes: Pre-reform cohorts are born 1967–1969, phase in-cohorts are born 1970–1972, and post-reform cohorts are born 1973–1976. Treatment (comparison) municipalities are above (below) the median in child care coverage growth from 1976 to 1979. Control variables are defined in Section 5. Standard deviations are in square brackets, and normalized differences are in curly brackets.

of the child care investment funds from other policies affecting child development, then our estimates will be biased downwards. Similarly, if municipalities expand in order to stimulate a particularly positive trend or if expansive municipalities also invest in other means of stimulating child development, then our estimates will be biased upwards. It is useful, therefore, to understand the determinants of the expansion across municipalities.

In Appendix A, we include a map of Norway, marking the treatment and comparison municipalities in Figure A1. The map shows that the municipalities are reasonably well spread out, covering urban and rural municipalities. In our baseline specification, five of the ten largest cities – by the number of children in our sample – are defined as treatment municipalities (Oslo, Bergen, Stavanger, Bærum and Fredrikstad), while the others are defined as comparison municipalities (Trondheim, Kristiansand, Tromsø, Skien and Drammen). Further, Table A1 in Appendix A display characteristics of the municipalities in the treatment and comparison area. There appears to be no substantial differences in terms of local government expenditure per capita, in total or on primary school in particular. This is most likely because of strict federal provisions for minimum standards of different local public services, and considerable ear-marked grants-in-aid from the central government. The same holds for local government income, consisting largely of grants-in-aid from the central government, local income taxes, and user fees. This comes as no

surprise, as the federal government determines the tax rate and the tax base of the income tax. Also, the federal government used equalization transfers to redistribute income from rich to poor municipalities, such that local differences in revenues are largely offset (Løken, 2009). Interestingly, there are no noticeable differences in the share of female voters between the municipalities of the treatment and comparison area, nor is there significant disparity in the socialist shares of voters. This conforms well to the fact that there was broad bipartisan support for child care expansion in Norway in the 1970s. Further, we do not find any substantial differences in population size or the population shares of neither 0 to 6 year olds, nor females of fecund age (19–35 or 36–55 years old).

There are, however, a couple of notable differences between treatment and comparison municipalities. Most importantly, the ratio of child care coverage to employment rate of mothers of 3-6 year olds prior to the reform, is substantially lower in treatment municipalities than in comparison municipalities. In treatment municipalities, there is on average more than four employed mothers for each child care place, while the same ratio is less than three-to-one in comparison municipalities. This conforms well to intuition, since federal subsidy rates were higher for municipalities with low child care coverage prior to the reform, but also because the local political pressure for expansion of formal care is likely to be stronger in areas where child care was severely rationed. We also see that two of the variables reflecting rurality indicate a small positive relationship with the child care expansion (average distance to zone center and ear marks per capita). This might be due to the discreteness of child care expansion; Establishing a typical child care institution increases the child care coverage rate more in smaller than in larger municipalities. In Norway, there was a very slow process of urbanization until the mid 1980s (Berg, 2005), which implies that rurality status is likely to be more or less constant during the period we consider, and should, therefore, be picked up by the municipality-specific fixed effects.²⁰

6 Empirical results

This section reports our main findings, before presenting results from a battery of specification checks.

6.1 Main results

Figure 5 draws the estimated effect from the threshold DD model, defined in equation (2). This figure shows the estimated effects per child care place on the probability of earnings

²⁰We have also regressed the change in the municipality’s child care coverage between 1976 and 1979 on the characteristics of the municipalities listed in Table A1. Consistent with the descriptive statistics, there is little evidence of systematic relationships between the child care expansion and most of these characteristics. Again, the most notable exception is the ratio of child care to maternal employment rate prior to the reform.

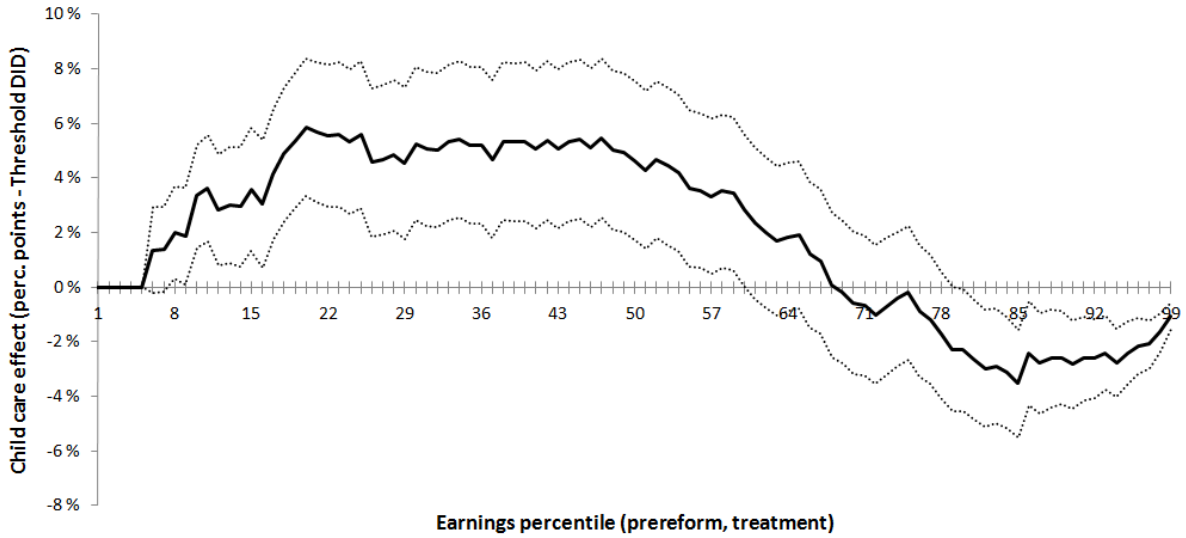


Figure 5: Threshold DD estimate and 90%-confidence interval of the effect of the child care reform on the probability of earning above pre-treatment percentiles 1–99, including covariates.

Notes: Estimations are based on OLS on equation (2) with controls listed in Table 3. The dependent variable is an indicator for earnings above the observed percentiles in treatment municipalities among pre-reform cohorts. Estimates and standard errors are corrected for intention to treat equal to .1785, see Section 4. The sample size is 498,956. Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

above pre-treatment percentiles, with a 90-percent two-sided confidence bound. The pre-treatment percentiles used in the estimations can be seen in the earnings distributions of Figure 4.

Figure 5 reveals major heterogeneity in the child care effects. The estimated effects are zero or positive for all percentiles until the 69th, before turning negative at the upper part of the distribution. This heterogeneity is consistent with our theoretical predictions: Children from disadvantaged families benefit the most from the child care reform, whereas it is less important or even detrimental for children from families with monetary and human capital to facilitate alternative arenas for child development of relatively high quality. While the negative effects at the top of the distribution are quite substantial with a decrease between 1 and 3.5 percentage points above the 77th percentile, the effects at the bottom and in the middle are generally larger: The distribution is lifted by 4.6 percentage points at the median, by 5.9 percentage points at the 20th percentile and no less than 2.8 percentage points between the 10th percentile and the 60th percentile. It is also clear that the estimated effects are fairly precise, with significant positive effects at every percentile between the 10th and the 60th, and significant negative effects from the 81st percentile (according to significance level of .05 for a two-sided test).

Having estimated the impact of the child care reform on the entire earnings distribution enables us to map out the associated counterfactual distribution that would have prevailed

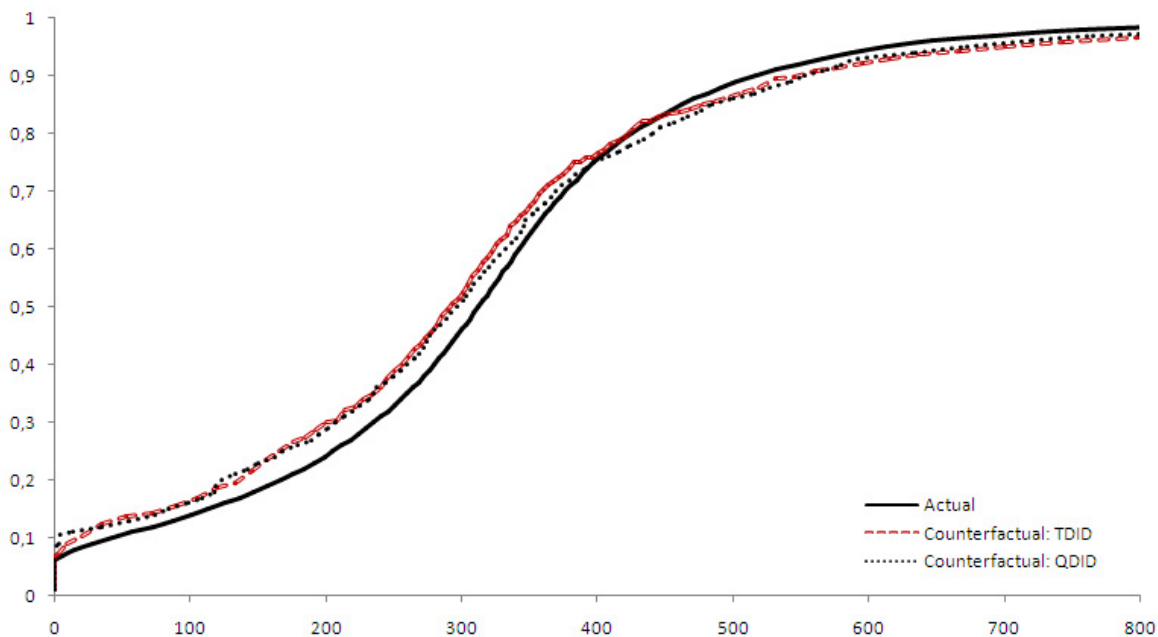


Figure 6: Actual earnings distribution (NOK 1,000) among post-reform cohorts from treatment municipalities, and counterfactual earnings distribution in the absence of the reform, estimated using the Quantile DD- and the Threshold DD-approach.

Notes: The difference between actual and counterfactual earnings distributions are based on estimates in Figure 5 and Figure B1. Post-reform cohorts are born 1973–1976. Treatment (comparison) municipalities are above (below) the median in child care coverage growth from 1976 to 1979.

in the absence of the reform. Figure 6 draws the actual and counterfactual earnings distributions observed among post-reform cohorts from treatment municipalities. The horizontal distance reflect the child care effects measured in earning levels at different quantiles, while the vertical distance displays the child care effects measured in population shares at different earnings thresholds.

The figure reveals that the counterfactual distribution lies above the actual distribution up to the intersection point at the 68th percentile, and after that below. It is also evident that the differences in population shares are largest below the median, whereas the differences in earnings levels are highest in the upper part of the distributions. However, the percentage change in earnings because of the reform is much larger at the lower part of the distribution than at the upper part of the distribution. For example, our findings suggest that earnings increase by more than 30 percent per child care place below the 21st percentile percentile and by more than 7 percent below the median.

Mean effects. When comparing the results from the threshold DD method with the estimated mean impact for the population as a whole, it is clear that mean impacts miss a lot. Panel A in Table 4 shows a mean impact of NOK $-3,194$, indicating that the child care reform was, on average, harmful to children’s earnings as adults. However, the mean impact is imprecisely estimated, and we cannot rule out a positive average effect

Table 4: Mean effects

	Estimate	SE	Mean	N
<i>Panel A. Full sample</i>				
Earnings	-3,194	7,790	322,893	498,956
<i>Panel B. Father's earnings</i>				
'- under median	24,594*	13,348	314,709	252,735
'- over median	-2,869	22,205	331,293	246,221
'- under 25th perc.	51,627**	21,288	303,935	126,738
'- over 75th perc.	34,948	50,133	336,638	121,946
<i>Panel C. Parental education</i>				
Mother, junior high or less	14,650*	8,168	311,375	397,110
'- more than junior high	-25,212	21,860	367,804	101,846
Father, junior high or less	15,960*	8,857	304,586	312,578
'- more than junior high	-20,410	14,758	353,595	186,378
<i>Panel D. Child sex</i>				
Girls	6,533	8,636	252,445	245,270
Boys	-14,238	12,793	391,004	253,686

Notes: Estimations are based on OLS on equation (1) with controls listed in Table 3 with earnings as the dependent variable. Estimates and standard errors are corrected for intention to treat equal to .1785, see Section 4. The sample size is 498,956. Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

on children's earnings. Nevertheless, our non-linear DD results suggest that we should interpret the mean impact estimate with caution, as it masks substantial heterogeneity in the child care effects.

The most common way to allow for such heterogeneity is to perform sub-sample analysis. To investigate whether the essential features of our findings from the non-linear DD method can be revealed by sub-sample analysis, we follow the previous literature closely in estimating equation (1) separately by child and parental characteristics. Panel B–D in Table 4 reports the sub-sample results.²¹

Panel B displays sub-sample results according to father's average earnings in the years the child were eligible for child care, that is when the child was between 3 and 6 years old.²²

²¹We have also performed sub-sample analysis by number of siblings, parents' age at birth, and mother's earnings. The sub-sample results by parental age suggest that children of younger parents benefitted slightly, while children of older parents are estimated to lose from the child care reform. These results conform well to the sub-sample results by parental education and father's earnings, which are positively associated with parental age. The results for mother's earnings and family size are close to zero, but very imprecise.

²²Arguably father's earnings is potentially endogenous to the reform. Ideally, we should therefore use father's earnings in the years prior to child care eligibility. Unfortunately, we do not have data on father's earnings for the pre-reform cohorts before the child is 3 years old. Also, Havnes and Mogstad (2009) show that the child care had little, if any, effect on maternal (part-time and full-time) employment. Instead of increasing mothers' labor supply, the new subsidized child care mostly crowds out informal child care

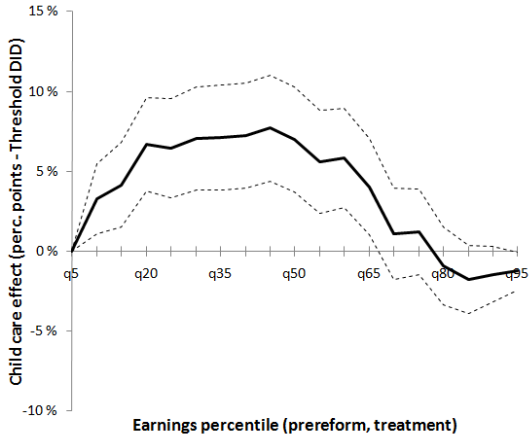
Conforming to the theoretical predictions, our result suggests different effects depending on fathers' earnings. Children of father's with above-median earnings suffer a loss from the reform estimated at NOK $-2,869$ per child care place. In contrast, children of fathers below the 25th percentile are estimated to benefit by NOK $51,627$ per child care place. Given the large body of evidence suggesting a strong intergenerational transmission in earnings, in particular from father to child (see e.g. Dearden et al., 1997), these sub-sample results indicate that the child care reform had an equalizing effect, which aligns well with the findings from the non-linear DD methods. The results in Panel C support this contention: Most of the benefits associated with universal child care relate to children of low educated parents, whereas children of high educated parents actually experience a loss from the child care reform.

Finally, Panel D splits the sample by the sex of the child. Although our theoretical framework makes no predictions about heterogeneity in the effects according to sex of the child, previous studies suggest that girls benefit the most from child care exposure (see e.g. Melhuish et al., 2008), and boys may actually lose (see e.g. Datta Gupta and Simonsen, 2007). Our sub-sample results echo these findings: While the earnings of girls exposed to the child care reform are estimated to increase by NOK $6,533$ per child care place, the earnings of boys decreased by NOK $-14,238$ per child care place. Given that females earn, in general, less than males, these sub-sample results also indicate that the child care reform had an equalizing effect on the earnings distribution.

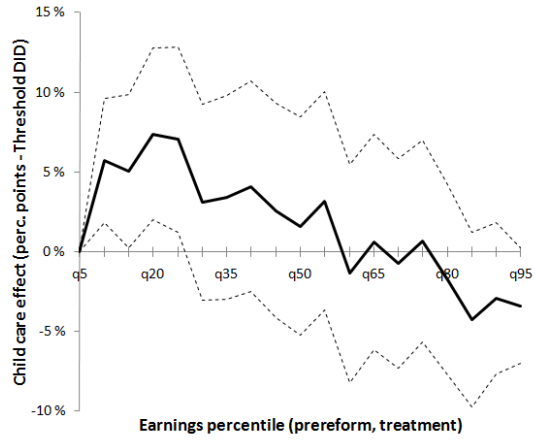
The general picture from Table 4 is that the sub-sample results reveal some heterogeneity. Moreover, the heterogeneous effects are broadly consistent with our theoretical predictions and the non-linear DD results. However, the sub-sample results are too imprecise to draw firm conclusions about heterogeneous effects, let alone judge whether child care was leveling the playing field or not. Further, the evidence from the non-linear DD methods of positive effects on most of the earnings distribution is not mirrored particularly well in the sub-sample analysis. As suggested by Bitler et al. (2006), estimating sub-group mean impacts may miss essential heterogeneity and, moreover, suffer from imprecision, if the intra-group variation in effects is relatively large compared to the inter-group variation. Figure 7 explores this issue, estimating the threshold DD method separately by subgroup.²³ The results reveal large intra-group variation in the child care effects: the within-group heterogeneity is characterized by sizable positive effects on the lower and the central parts of the earnings distribution.

arrangements. Since most fathers were already working, we would expect the impact to be even smaller on them.

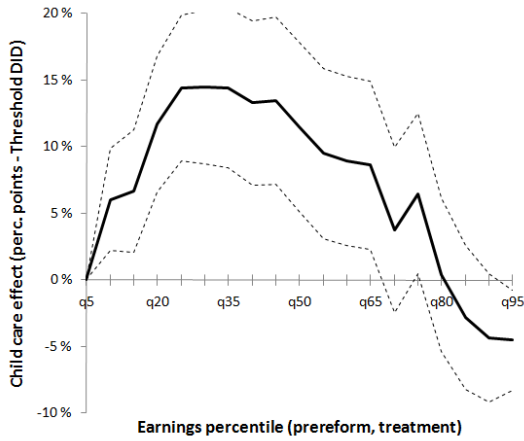
²³Estimations are performed on 20 quantiles for computational simplicity.



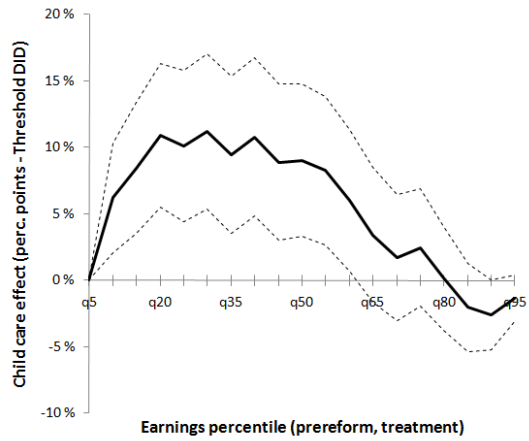
(a) Mom low educated



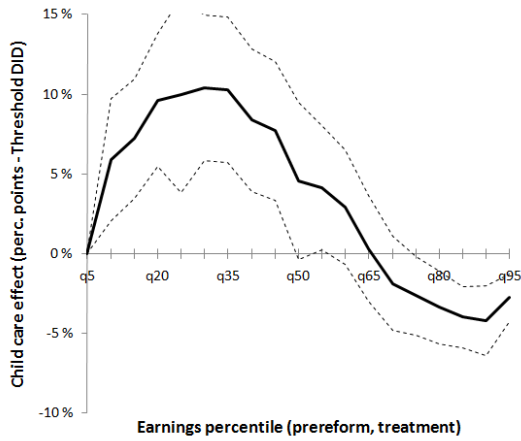
(b) Mom high educated



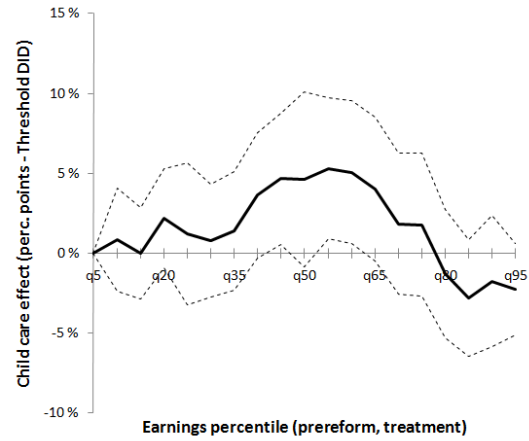
(c) Dad earnings below median



(d) Dad earnings above median



(e) Girls



(f) Boys

Figure 7: Threshold DD-estimates by subsample.

Notes: Estimations are based on OLS on equation (2) with controls listed in Table 3. The dependent variable is an indicator for earnings above the observed percentiles in treatment municipalities among pre-reform cohorts. Estimates and standard errors are corrected for intention to treat equal to .1785, see Section 4. Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

6.2 Robustness analysis

This section reports results from our robustness analysis. In general, the results from the large number of specification checks align well with our main findings from the threshold DD method, showing positive effects at the lower and central parts and negative effects at the upper parts of the earnings distribution. Also the mean impact estimates are, by and large, quite similar across the specifications – though suffering from imprecision, just like the mean impacts estimation in our main results. As the focus of our study is on distributional effects, we will concentrate our discussion below on the robustness results from the threshold DD method, reported in Table 5. Since the threshold DD method is computationally burdensome, the specification checks are only performed at certain percentiles, covering different parts of the earnings distribution. For brevity, we do not report robustness results for the mean effects.

Time trend. The DD approach controls for unobserved differences between children born in different years as well as between children from treatment and comparison municipalities. The identifying assumption is that the change in earnings for 3 to 6 year olds before and after the reform would have been the same in the treatment municipalities as in the comparison municipalities, in the absence of the reform. A concern is that our estimated effects may reflect differential time trends between the treatment and comparison municipalities, rather than true policy impact. To investigate the common trend assumption, we perform a number of specification checks.

Our first robustness check is to estimate equation (1) with and without the set of controls X . This investigates whether the time trend in children’s outcomes differs by, say, parent’s education, while there are systematic changes over time in parental education in treatment municipalities compared to comparison municipalities. Substantial changes over time in the differences in the observable characteristics of the two groups may suggest unobserved compositional changes, calling our empirical strategy into question. Column (2) reports estimation results without controls, whereas Column (1) repeats our baseline results for comparison. In line with the results from the baseline specification, the threshold DD estimates without controls suggest positive effects in the lower and middle parts of the earnings distribution and negative effects at the upper part.

Next, we perform two placebo tests. In the first placebo test, we pretend that the child care expansion took place in the pre-reform period. Specifically, we add interaction terms between treatment status and cohort dummies for children born in 1968 and 1969 (with the 1967-cohort as the omitted category) to equation (2). The placebo test turns on the estimated treatment effects for cohorts born in 1968 and 1969, relative to 1967. Column (3) shows that none of the placebo reform effects for cohorts born in 1969 is significant at conventional levels. The same is true for placebo reform effects for the 1968 cohort, which are omitted for brevity.

Although it is reassuring to find that the trends are not systematically deviating in the pre-reform period, we may worry about breaks in the underlying trends coinciding with the reform. If we could find a variable that is strongly correlated with earnings, but unlikely to be affected by the child care reform, then we could tackle this by performing a placebo test within the reform period. Our second placebo test therefore exploits variation in adult height. Height should be a promising candidate for two reasons. First, a large number of twin and adoptive studies have shown that genetic factors are the overwhelming determinant of variation in height within developed countries. For example, Silventoinen et al. (2003) report heritability estimates around 0.9 for Norwegian males born between 1967 and 1978, implying that within this population about 90% of the variance in adult height can be accounted for by the variance of genes.

Second, it has long been recognized that taller adults have, on average, higher earnings. This also holds true in our sample where the correlation between height and earnings is strong and significant: For example, a one standard deviation increase in height is associated with an increase in earnings of almost USD 3,400 (over NOK 22,000). Case and Paxson (2008) offer an explanation: On average, taller people take higher education and earn more because they are smarter. As early as age 3 – before child care has had a chance to play a role – and throughout childhood, they find that taller children perform significantly better on cognitive tests. Moreover, they demonstrate that the correlation between height in childhood and adulthood is very high, so that tall children are much more likely to become tall adults.

Significant effects of the child care reform on children’s adult height would therefore raise concern that effects on other outcomes reflect omitted variables bias, like unobserved heterogeneity in innate ability, rather than true policy impacts. It is therefore re-assuring that the estimates reported in Column (4) are close to zero and insignificant, suggesting no effect of the child care reform on children’s adult height.

To make sure that results are not driven by secular changes between urban and rural areas coinciding with the reform, we further drop the three big cities (Oslo, Bergen, and Trondheim) from our analysis. Column (5) reports estimation results excluding these cities. The fact that our estimates vary little between the specifications increases our confidence in the empirical strategy.

Although municipality-specific fixed effects picks up the direct effects of pre-determined factors of the municipalities, like differences in rationing of formal child care prior to the reform, we may worry about the determinants of the child care expansion being systematically related to underlying trends in child outcome. To address this concern, we follow Duflo (2001) in allowing differential inter-cohort time trends across municipalities in the DD estimation. Specifically, we interact the cohort fixed effects with a large set of pre-reform municipality characteristics, such as male and female education and labor supply, primary school expenditures, and population density. These characteristics are

listed in Table A1 in the Appendix, and discussed in Section 5. In doing so, we allow for different underlying trends in children’s potential outcomes, depending on the pre-reform characteristics of the municipality. Column (9) reports estimates from this specification, which conform well to the results from the baseline specification.

Selective migration. Although location decisions based on unobservable characteristics may affect our estimates, the direction of the bias is not obvious.²⁴ On the one hand, education-oriented or labor market attached parents may be more likely to move to municipalities with high child care coverage rate. On the other hand, parents with children who need special attention or supervision may be more inclined to move to municipalities with good access to child care. Though recent empirical work finds little support for Tiebout sorting across states or municipalities according to public good provision like school quality,²⁵ we take several steps to avoid that selective migration of families into treatment and comparison municipalities confounds our results.

To address the concern for in-migration induced by the reform, we excluded in our main analysis children from families that move between treatment and comparison municipalities during the expansion period. In addition, we control for relocation between municipalities within the treatment/comparison area; We have also performed all estimations excluding families that relocate, and the results are unchanged. However, one could argue that even the sample of stayers is selective, as out-migration could be endogenous to the child care expansion. To address this issue, we follow Hægeland et al. (2008) in using children’s municipality of birth to determine whether they belong to treatment or comparison municipalities. Column (6) shows that the effects of the child care expansion is fairly robust to using municipality of birth to determine treatment status. This finding conforms well with the results from Hægeland et al. (2008), which suggest that school quality matters little, if anything, for location decisions in Norway.

Alternative treatment definitions. In our baseline specification, we define the treatment and comparison areas by ordering municipalities according to the increase in child care coverage rate in the period 1976–1979, and then separating them at the median. Below, we make sure that our results are not artifacts of this choice of treatment definition.

In Column (7), we use the same expansion period, 1976–1979, but divide the sample at the 33rd and 67th percentiles of child care growth. Municipalities below the lower threshold are in the comparison group, while those above the upper threshold are in the treatment group. Children from municipalities in between the thresholds are excluded from the sample used for estimation. In Column (8), we define the treatment and comparison according to the median child care growth, but alter the expansion period to

²⁴Note that families living on the municipal borders could not take advantage of the child care expansion in neighboring municipalities without relocating, since eligibility was based on municipality of residency.

²⁵See e.g. Rhode and Strumpf (2003) who find little support for Tiebout sorting across municipalities and counties using about 150 years of data.

Table 1: Robustness: Threshold DD

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Baseline	No	Placebo	Height	No	Selective	1976-79	1977-79	Flexible	Linear in
		covariates			cities	migration	33rd/67th	50th/50th	trends	child care
5th percentile	0	0	0	-0.0004 (0.0025)	0	0	0	0	0	0
10th percentile	0.0335*** (0.0114)	0.0288** (0.0114)	0.0054 (0.0201)	-0.0031 (0.0032)	0.0335*** (0.0114)	0.0189 (0.0117)	0.0166 (0.0160)	0.0288** (0.0114)	0.0256** (0.0122)	0.0120 (0.0077)
25th percentile	0.0558*** (0.0164)	0.0599*** (0.0169)	-0.0056 (0.0282)	-0.0001 (0.0045)	0.0558*** (0.0164)	0.0459*** (0.0168)	0.0654*** (0.0236)	0.0358** (0.0164)	0.0527*** (0.0178)	0.0374*** (0.0114)
50th percentile	0.0463*** (0.0176)	0.0621*** (0.0192)	0.0360 (0.0307)	-0.0001 (0.0048)	0.0463*** (0.0176)	0.0273 (0.0181)	0.0752*** (0.0255)	0.0373** (0.0176)	0.0427** (0.0192)	0.0250** (0.0122)
75th percentile	-0.0020 (0.0148)	0.0125 (0.0158)	0.0433 (0.0268)	-0.0019 (0.0041)	-0.0020 (0.0148)	-0.0072 (0.0152)	0.0162 (0.0214)	-0.0031 (0.0148)	0.0080 (0.0161)	-0.0119 (0.0101)
90th percentile	-0.0283*** (0.0099)	-0.0215** (0.0101)	0.0075 (0.0190)	-0.0006 (0.0028)	-0.0283*** (0.0099)	-0.0363*** (0.0102)	-0.0355** (0.0144)	-0.0297*** (0.0099)	-0.0162 (0.0106)	-0.0301*** (0.0066)
95th percentile	-0.0241*** (0.0070)	-0.0202*** (0.0071)	0.0069 (0.0139)	-0.0013 (0.0019)	-0.0241*** (0.0070)	-0.0231*** (0.0072)	-0.0227** (0.0101)	-0.0204*** (0.0070)	-0.0157** (0.0074)	-0.0212*** (0.0046)
ITT	0.1785	0.1785	0.1785	0.1785	0.1785	0.1785	0.2946	0.1171	0.1785	-
No. of children	498,956	498,956	498,956	246,516	498,956	483,327	259,658	498,956	451,530	498,956

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Each entry reports the treatment effect corrected for intention to treat (ITT), see Section 4. In columns 1-9, estimations are based on OLS on equation (2) with controls listed in Table 3. The dependent variable is an indicator for earnings above the observed percentiles in treatment municipalities among pre-reform cohorts, except in column 3 where the dependent variable is defined from the distribution of height. Column 1 repeats our baseline estimates. In column 2, we report results when excluding our set of controls. In column 3, the treatment effect is estimated for the pre-reform cohort born in 1969. In column 4, the dependent variable is defined from the distribution of height among males at age 18. In column 5, the three largest cities are dropped. In column 6, treatment is defined by the child's birth municipality (children with missing values dropped). In column 7, treatment municipalities are above the 67th percentile in child care coverage growth, while comparison municipalities are below the 33rd. In column 8, the expansion period is 1977-1979. In column 9, the cohort fixed effects are interacted with pre-reform municipality characteristics listed in Table A1. In column 10, estimates are based on equation (3). Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

1977–1979. To be consistent with this new definition of the expansion period, the 1970 cohort is now defined as a pre-reform instead of a phase-in cohort. The results in Columns (7) and (8) echo our main results.

To sidestep the issue of how to draw the line between treatment and comparison municipalities, we follow Berlinski et al. (2009) in regressing children’s earnings on child care coverage rate in each municipality, controlling for cohort and municipality fixed-effects, as well as a set of controls. This non-linear DD specification restricts the marginal effects of additional child care slots to be constant, and can be defined as

$$\Pr(Y_{ijt} \geq y) = \delta_t^y + \zeta^y CC_{it} + X'_{ijt} \varphi^y + \epsilon_{ijt}^y \quad (3)$$

where CC_{it} is the average child care rate in the municipality of child i from the year t when the child turns 3 years old until, but not including, year $t + 4$ when he or she turns 7 and starts primary school. In line with the results from the baseline specification, the findings reported in Column (10) suggest positive effects in the lower and middle parts of the earnings distribution and negative effects at the upper part.

Definition of earnings. A concern is that individuals’ earnings may fluctuate over time due to transitory shocks, as well as institutional factors such as accounting and tax rules. Columns 2 and 3 in Table 6 report results from the threshold DD method, using annual earnings in 2007 and average earnings between 2006 and 2007 as the dependent variable. The estimates confirm the picture of positive effects at the lower and central parts and negative effects at the upper parts of the earnings distribution.

Quantile DD results. Our final specification check employs the quantile DD method to estimate the distributional effects of the child care reform. Even though the identifying assumption is different, the results from the quantile DD method reported in the panels of Figure B1 in Appendix B are quite similar to the findings from the threshold DD method.²⁶ Specifically, the quantile treatment effects are positive at most percentiles, before turning negative at the uppermost part of the distribution. Figure 6 draws the actual and counterfactual earnings distributions among post-reform cohorts from treatment municipalities, according to the quantile DD method and the threshold DD method, showing a very similar picture.

It should be noted that panel (b) in Figure B1 displays major fanning-out of the confidence interval at the upper end of the distribution. This is most likely due to the use of inverse propensity score-weighting to handle covariates. A drawback with this procedure is that the results may become rather unstable as the propensity tends towards 0

²⁶It should be noted that quantile DD results are reported in NOK, whereas threshold DD results are reported in percentage points. Also, the earnings levels at which threshold DD-estimation is performed are defined from the pre-reform earnings distribution in the treatment group, which is not a cardinal scale. Specifically, the distance between quantiles is usually increasing as we move up in the distribution, see Figure 4. For a direct comparison of results from quantile and threshold DD, see Figure 6

Table 6: Alternative income measures: TDID

	(1) Baseline	(2) Earnings 2006–2007	(3) Earnings 2007
5th percentile	0 –	0 –	0 –
10th percentile	0.0335*** (0.0114)	0.0266** (0.0113)	0.0279** (0.0112)
25th percentile	0.0558*** (0.0164)	0.0460*** (0.0163)	0.0423*** (0.0162)
50th percentile	0.0463*** (0.0176)	0.0392** (0.0176)	0.0474*** (0.0176)
75th percentile	-0.0020 (0.0148)	-0.0171 (0.0149)	-0.0123 (0.0149)
90th percentile	-0.0283*** (0.0099)	-0.0217** (0.00999)	-0.0228** (0.0101)
95th percentile	-0.0241*** (0.00699)	-0.0248*** (0.00705)	-0.0209*** (0.00715)
No. of children	498,956	495,523	495,523

Notes: Estimations are based on OLS on equation (2) with controls listed in Table 3. The dependent variable is an indicator for earnings above the observed percentiles in treatment municipalities among pre-reform cohorts. Estimates and standard errors are corrected for intention to treat equal to .1785, see Section 4. Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

or 1.²⁷ Panel (a) therefore reports results from a quantile DD-estimation without covariates, which can be compared to the results from the threshold DD-estimation without covariates reported in Figure A2 in Appendix B. While the results from the threshold DD method is robust to inclusion and exclusion of covariates, the results from the quantile DD method are sensitive to this choice at the upper end of the distribution. Specifically, the estimated effects are very similar until the upper most part of the distribution, where the quantile DD estimates with covariates in panel (b) do not shoot down, as when covariates are excluded in panel (a). The results in panel (b) also never turn significantly negative.

7 Conclusion

There is a heated debate in the US and Canada, as well as in many European countries, about a move towards subsidized, universally accessible child care or preschool. Advocates of a universal child care system offer a two-fold argument: Universal child care is claimed to both facilitate children’s long-run development, and level the playing field by benefiting in particular disadvantaged children. Therefore, a critical element in evaluating universal child care systems is measuring the impact on child development in a way that allows the

²⁷See Frölich (2004) for a discussion and Monte Carlo-evidence.

effects to vary systematically across the outcome distribution.

This paper contributes by providing first evidence on the distributional effects of universal child care on child outcome. Using non-linear DD methods, we investigate how the introduction of subsidized, universally accessible child care in Norway affected the earnings distribution of exposed children as adults. Our robust findings lead to three conclusions. First, mean impacts miss a lot, concealing major heterogeneity: While child care had a small and insignificant mean impact, effects were positive over the bulk of the earnings distribution, and sizable below the median. This is an important observation since previous empirical studies of universal child care have focused on mean impacts. Second, the estimated heterogeneity in child care effects is consistent with a simple theoretical framework where parents trade off own consumption with investment in child quality. Third, the essential features of our empirical findings could not have been revealed using mean impact analysis on typically defined subgroups. This is because the intragroup variation in the child care effects is relatively large compared to the intergroup variation in mean impacts.

In sum, our study shows that non-linear DD methods can play a useful role in assessing policy changes, when only non-experimental data is available and theory predicts heterogeneous treatment effects. The results from our study are also of interest from a policy perspective: Although the child care reform failed to improve, on average, children's earnings prospects, it did succeed in leveling the playing field. This illustrates how the conclusions drawn from policy evaluations of universal child care policies might depend crucially on the trade-off between the mean and (in)equality in the distribution of children's outcomes.

References

- ALMOND, D., AND J. CURRIE (2010): “Human Capital Development Before Age Five,” NBER Working Papers 15827, National Bureau of Economic Research, Inc.
- ANGRIST, J. D. (2001): “Estimations of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice,” *Journal of Business & Economic Statistics*, 19(1), 2–16.
- ATHEY, S., AND G. W. IMBENS (2006): “Identification and Inference in Nonlinear Difference-in-Differences Models,” *Econometrica*, 74(2), 431–497.
- ATKINSON, A. B., L. RAINWATER, AND T. M. SMEEDING (1995): *Income distribution in OECD countries : evidence from the Luxembourg Income Study*. OECD Publications and Information Center, Paris.
- BAKER, M., J. GRUBER, AND K. MILLIGAN (2008): “Universal Child Care, Maternal Labor Supply, and Family Well-Being,” *Journal of Political Economy*, 116(4), 709–745.
- BARNETT, W. S. (1995): “Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes,” *The Future of Children*, 5(3), 25–50.
- BECKER, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education*. Univ of Chicago Press, Boston, MA.
- BERLINSKI, S., S. GALIANI, AND P. GERTLER (2009): “The effect of pre-primary education on primary school performance,” *Journal of Public Economics*, 93(1-2), 219–234.
- BERLINSKI, S., S. GALIANI, AND M. MANACORDA (2008): “Giving children a better start: Preschool attendance and school-age profiles,” *Journal of Public Economics*, 92(5-6), 1416–1440.
- BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): “How Much Should We Trust Differences-in-Differences Estimates?,” *The Quarterly Journal of Economics*, 119(1), 249–275.
- BITLER, M. P., J. B. GELBACH, AND H. W. HOYNES (2006): “What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments,” *American Economic Review*, 96(4), 988–1012.
- BOHLMARK, A., AND M. J. LINDQUIST (2006): “Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden,” *Journal of Labor Economics*, 24(4), 879–900.

- BUCHINSKY, M. (1995): “Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study,” *Journal of Econometrics*, 68(2), 303–338.
- CARNEIRO, P., AND J. J. HECKMAN (2004): “Inequality in America: What Role for Human Capital Policies?,” in *Human Capital Policy*, ed. by J. J. Heckman, and A. B. Krueger. The MIT Press.
- CASCIO, E. U. (2009): “Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools,” NBER Working Papers 14951, National Bureau of Economic Research, Inc.
- CASE, A., AND C. PAXSON (2008): “Stature and Status: Height, Ability, and Labor Market Outcomes,” *Journal of Political Economy*, 116(3), 499–532.
- CHEN, Y., AND L.-A. ZHOU (2007): “The long-term health and economic consequences of the 1959-1961 famine in China,” *Journal of Health Economics*, 26(4), 659–681.
- CUNHA, F., AND J. J. HECKMAN (2008): “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation,” *Journal of Human Resources*, 43(4), 738–782.
- CURRIE, J. (2001): “Early Childhood Education Programs,” *Journal of Economic Perspectives*, 15(2), 213–238.
- DATTA-GUPTA, N., AND M. SIMONSEN (2007): “Non-cognitive Child Outcomes and Universal High Quality Child Care,” IZA Discussion Papers 3188, Institute for the Study of Labor (IZA).
- DEARDEN, L., S. MACHIN, AND H. REED (1997): “Intergenerational Mobility in Britain,” *The Economic Journal*, 107(440), 47–66.
- DUFLO, E. (2001): “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment,” *American Economic Review*, 91(4), 795–813.
- EIDE, M., N. ØYEN, R. SKJÆRVEN, S. NILSEN, T. BJERKEDAL, AND G. TELL (2005): “Size at Birth and Gestational Age as Predictors of Adult Height and Weight,” *Epidemiology*, 16, 175.
- EU (2002): “Presidency Conclusions,” Barcelona European Council 15 and 16 March 2002, Barcelona.
- FIRPO, S. (2007): “Efficient Semiparametric Estimation of Quantile Treatment Effects,” *Econometrica*, 75(1), 259–276.

- FIRPO, S., N. M. FORTIN, AND T. LEMIEUX (2009): “Unconditional Quantile Regressions,” *Econometrica*, 77, 953–973.
- FITZPATRICK, M. D. (2008): “Starting School at Four: The Effect of Universal Pre-Kindergarten on Children’s Academic Achievement,” *B.E. Journal of Economic Analysis and Policy*, 8(1), 1–38.
- FRLICH, M. (2004): “Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators,” *The Review of Economics and Statistics*, 86(1), 77–90.
- GAVIRIA, A. (2002): “Intergenerational mobility, sibling inequality and borrowing constraints,” *Economics of Education Review*, 21(4), 331–340.
- GORMLEY, WILLIAM T., J., AND T. GAYER (2005): “Promoting School Readiness in Oklahoma: An Evaluation of Tulsa’s Pre-K Program,” *Journal of Human Resources*, 40(3), 533–558.
- GULBRANDSEN, L., J. A. LEA, AND S. STOKKE (1982): “Barnetilsyn hos smbarnsfamilier,” INAS notat 82:13, Institutt for anvendt sosialvitenskapelig forskning, Oslo.
- HAIDER, S., AND G. SOLON (2006): “Life-Cycle Variation in the Association between Current and Lifetime Earnings,” *American Economic Review*, 96(4), 1308–1320.
- HAVNES, T., AND M. MOGSTAD (2009): “No Child Left Behind. Universal Child Care and Children’s Long-Run Outcomes,” Discussion Papers 582, Statistics Norway.
- HECKMAN, J. J., J. SMITH, AND N. CLEMENTS (1997): “Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts,” *Review of Economic Studies*, 64(4), 487–535.
- HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,” *Journal of Labor Economics*, 24(3), 411–482.
- HERBST, C. M., AND E. TEKIN (2008): “Child Care Subsidies and Child Development,” IZA Discussion Papers 3836, Institute for the Study of Labor (IZA).
- HGELAND, T., O. RAAUM, AND K. G. SALVANES (2005): “Pupil Achievement, School Resources and Family Background,” IZA Discussion Papers 1459, Institute for the Study of Labor (IZA).
- IMBENS, G., AND J. WOOLDRIDGE (2009): “Recent Developments in the Econometrics of Program Evaluation,” *Journal of Economic Literature*, 47(1), 5–86.

- KAROLY, L. A., M. R. KILBURN, AND J. S. CANNON (2005): *Early Childhood Interventions: Proven Results, Future Promise*. RAND Corporation, Santa Monica, CA.
- KNUDSEN, E. I., J. J. HECKMAN, J. L. CAMERON, AND J. P. SHONKOFF (2006): “Economic, neurobiological, and behavioral perspectives on building America’s future workforce,” *Proceedings of the National Academy of Sciences*, 103(27), 10155–10162.
- KOENKER, R. (2005): *Quantile Regression*, Econometric Society Monographs. Cambridge University Press, Cambridge, MA.
- LEIRA, A. (1992): *Welfare States and Working Mothers*. Cambridge University Press.
- LKEN, K. V. (2009): “Family income and children’s education: Using the Norwegian oil boom as a natural experiment,” *Labour Economics*, In Press.
- LOEB, S., M. BRIDGES, D. BASSOK, B. FULLER, AND R. W. RUMBERGER (2007): “How much is too much? The influence of preschool centers on children’s social and cognitive development,” *Economics of Education Review*, 26(1), 52 – 66.
- MAGNUSON, K. A., C. RUHM, AND J. WALDFOGEL (2007): “Does prekindergarten improve school preparation and performance?,” *Economics of Education Review*, 26(1), 33–51.
- MELHUISH, E., K. SYLVA, P. SAMMONS, I. SIRAJ-BLATCHFORD, B. TAGGART, M. B. PHAN, AND A. MALIN (2008): “The Early Years: Preschool Influences on Mathematics Achievement,” *Science*, 321(5893), 1161–1162.
- MEYER, B. D., W. K. VISCUSI, AND D. L. DURBIN (1995): “Workers’ Compensation and Injury Duration: Evidence from a Natural Experiment,” *American Economic Review*, 85(3), 322–40.
- NOU (1972:39): “Preschools,” White paper, Ministry of children’s and family affairs.
- OECD (2004): “Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries,” Mem background paper, OECD Economics Department, Barcelona.
- POTERBA, J. M., S. F. VENTI, AND D. A. WISE (1995): “Do 401(k) contributions crowd out other personal saving?,” *Journal of Public Economics*, 58(1), 1–32.
- RHODE, P. W., AND K. S. STRUMPF (2003): “Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990,” *American Economic Review*, 93(5), 1648–1677.

- SHONKOFF, J. P., AND D. A. PHILLIPS (2000): *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academy Press.
- SILVENTOINEN, K., S. SAMMALISTO, M. PEROLA, D. I. BOOMSMA, B. K. CORNES, C. DAVIS, L. DUNKEL, M. DE LANGE, J. R. HARRIS, J. V. HJELMBORG, M. LUCIANO, N. G. MARTIN, J. MORTENSEN, L. NISTICO, N. L. PEDERSEN, A. SKYTTHE, T. D. SPECTOR, M. A. STAZI, G. WILLEMSSEN, AND J. KAPRIO (2003): “Heritability of Adult Body Height: A Comparative Study of Twin Cohorts in Eight Countries,” *Twin Research*, 6, 399–408(10).
- TELHAUG, A. O., O. A. MEDIS, AND P. AASEN (2006): “The Nordic Model in Education: Education as part of the political system in the last 50 years,” *Scandinavian Journal of Educational Research*, 50(3), 245 – 283.
- THE NORWEGIAN MINISTRY OF CHILDREN AND FAMILY AFFAIRS (1998): “OECD Thematic Review of Early Childhood Education and Care Policy: Background Report from Norway,” Report, Ministry of children’s and family affairs.
- VOLCKMAR, N. (2008): “Knowledge and Solidarity: The Norwegian social-democratic school project in a period of change, 1945-2000,” *Scandinavian Journal of Educational Research*, 52(1), 1–15.

Appendices

Appendix A



Figure A1: Geographic location of treatment (white) and comparison (dark) municipalities

Table A1: Descriptive statistics for treatment and comparison municipalities in 1976

	Treatment		Comparison	
Child care/maternal employment rate	0.2471	[0.4596]	0.3542	[0.6213]
Child care coverage rate	0.0534	[0.0899]	0.0695	[0.0968]
Years of education, males	9.2256	[0.5514]	9.2174	[0.4699]
–, females	8.8198	[0.3820]	8.7672	[0.3313]
Earnings, males	0.3917	[0.0762]	0.4081	[0.0734]
–, females	0.1080	[0.0349]	0.1121	[0.0321]
Labor force part., males	0.8324	[0.0591]	0.8367	[0.0665]
–, females	0.2919	[0.0844]	0.2997	[0.0813]
–, mothers of 3–6 year olds	0.1903	[0.0753]	0.1953	[0.0710]
Expenditure (2006-USD/capita)				
Total	959.65	[291.05]	909.72	[229.55]
Primary school	241.78	[107.40]	223.40	[90.83]
Revenue (2006-USD/capita)				
Ear marks, total	569.54	[217.95]	546.49	[174.11]
–, primary school	87.78	[33.31]	87.45	[36.12]
Fees, total	124.65	[90.03]	105.02	[63.80]
–, primary school	0.86	[1.38]	0.97	[1.80]
Taxes	379.00	[105.00]	379.09	[114.75]
Geography				
In densely populated areas	0.4049	[0.2915]	0.4827	[0.2979]
Ave. distance to zone center	0.8876	[0.7789]	0.7732	[0.6788]
– to neighboring center	3.7768	[2.6130]	3.4297	[2.8039]
Population				
Total	9846	[36400]	9476	[13267]
Married	0.4664	[0.0277]	0.4618	[0.0346]
Divorced	0.0144	[0.0081]	0.0153	[0.0080]
Immigrant	0.0098	[0.0096]	0.0095	[0.0086]
0 to 6 years old	0.1077	[0.0177]	0.1141	[0.0177]
7 to 10 years old	0.0673	[0.0099]	0.0708	[0.0097]
11 to 18 years old	0.1293	[0.0127]	0.1314	[0.0133]
Females: 19 to 35 years old	0.1021	[0.0187]	0.1082	[0.0170]
–: 36 to 55 years old	0.1027	[0.0101]	0.1019	[0.0095]
Males: 19 to 35 years old	0.1175	[0.0152]	0.1227	[0.0137]
–: 36 to 55 years old	0.1096	[0.0091]	0.1077	[0.0092]
Politics				
Registered voters	6480	[26654]	5863	[8488]
–, female	0.4896	[0.0169]	0.4928	[0.0167]
Election participation	0.7243	[0.0587]	0.7093	[0.0563]
–, females	0.7094	[0.0666]	0.6962	[0.0632]
Female elected representatives	0.1521	[0.0807]	0.1394	[0.0635]
Socialist vote share	0.3864	[0.1654]	0.4031	[0.1651]
Socialist mayor	0.3140	[0.4652]	0.3671	[0.4832]
Female mayor	0.0097	[0.0981]	0.0145	[0.1198]

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (one-tailed)

Notes: Columns 1–4 report means and standard deviations across treatment and comparison municipalities, not weighted by population size. Earnings denotes pensionable income in NOK 100,000. Socialist parties are defined as RV, SV and DNA. Densely populated areas are contiguous zones of at least 200 people where the distance between houses is generally less than 50 meters (400 meters if separated by e.g. parks, rivers or industrial zones). Average distance to zone center is the mean predicted travel distance in km from a citizen's home to the most populous census area in a contiguous zone of more than 2,000 people within the municipality. Distance to neighboring center is the mean travel distance from the center of a census area to the closest center in another census area within the same economic zone. Standard deviations are in brackets.

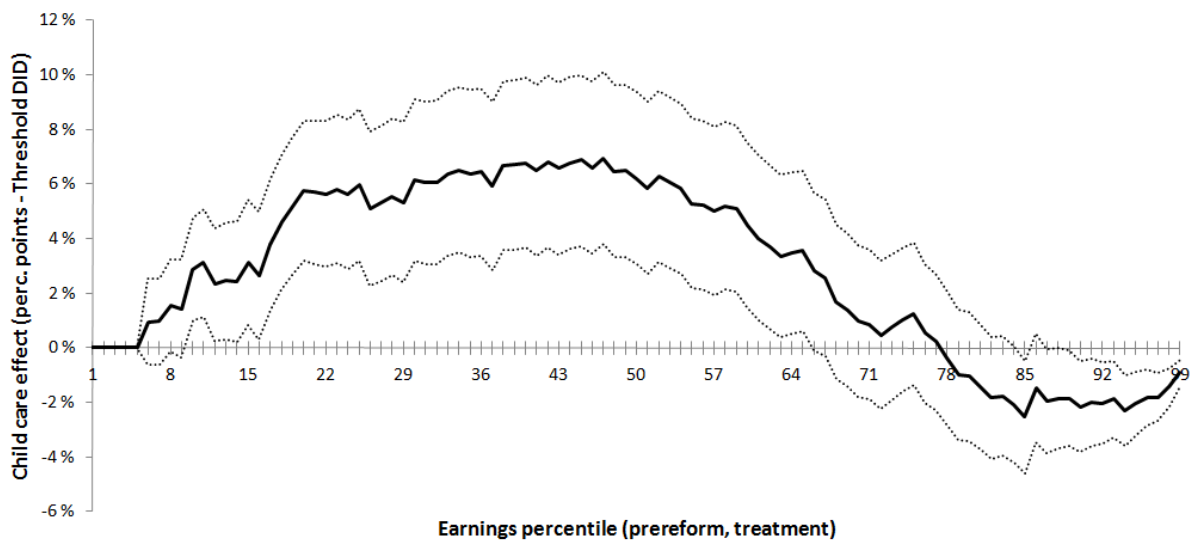


Figure A2: Threshold DD estimates and 90%-confidence interval of the effect of the child care reform on the probability of earning above pre-treatment percentiles 1–99, without covariates.

Notes: Estimations are based on OLS on equation (2) without covariates. The dependent variable is an indicator for earnings above the observed percentiles in treatment municipalities among pre-reform cohorts. Estimates and standard errors are corrected for intention to treat equal to .1785, see Section 4. The sample size is 498,956. Asymptotic standard errors are robust to within family clustering and heteroscedasticity.

Appendix B: Quantile DD

To handle covariates in the quantile DD model, we follow Firpo (2007) in using an inverse propensity score method. Specifically, the weights, ω , are given by

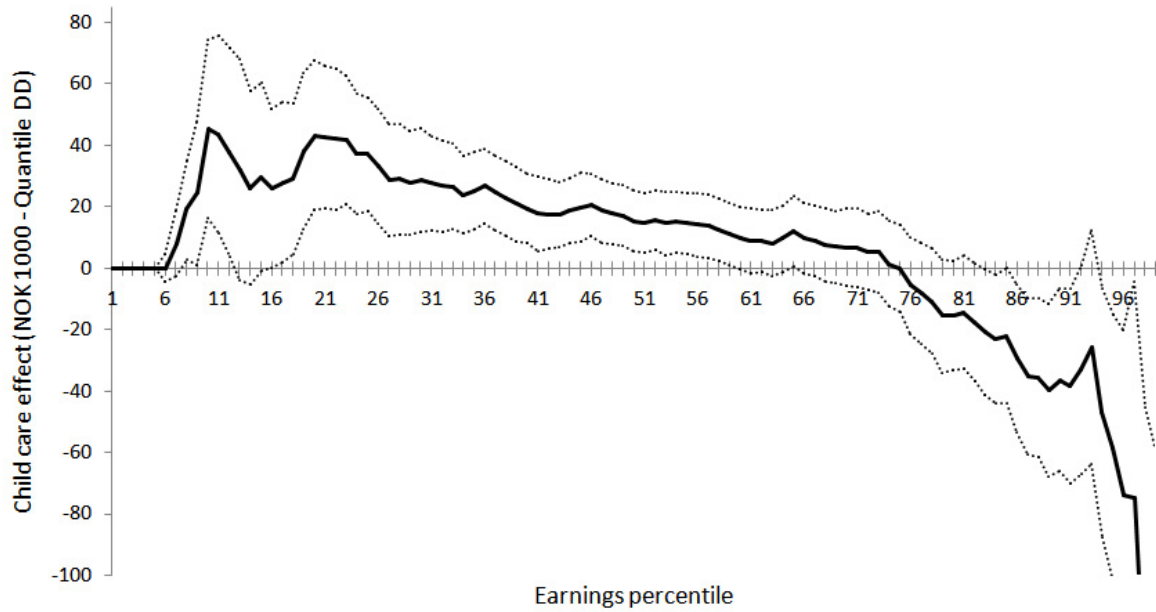
$$\hat{\omega}(X_i; Treat_i) = \frac{Treat_i}{N\hat{p}(X_i)} + \frac{1 - Treat_i}{N(1 - \hat{p}(X_i))} \quad (B1)$$

where N is the size of the estimation sample, and $\hat{p}(X_i)$ is the estimated propensity score for individual i from a logit regression. Let quantile q - in group g at time t in the weighted sample be denoted $\hat{F}_{gt}^{-1}(q; X)$. On the basis of the weighted sample, the quantile DD is estimated as

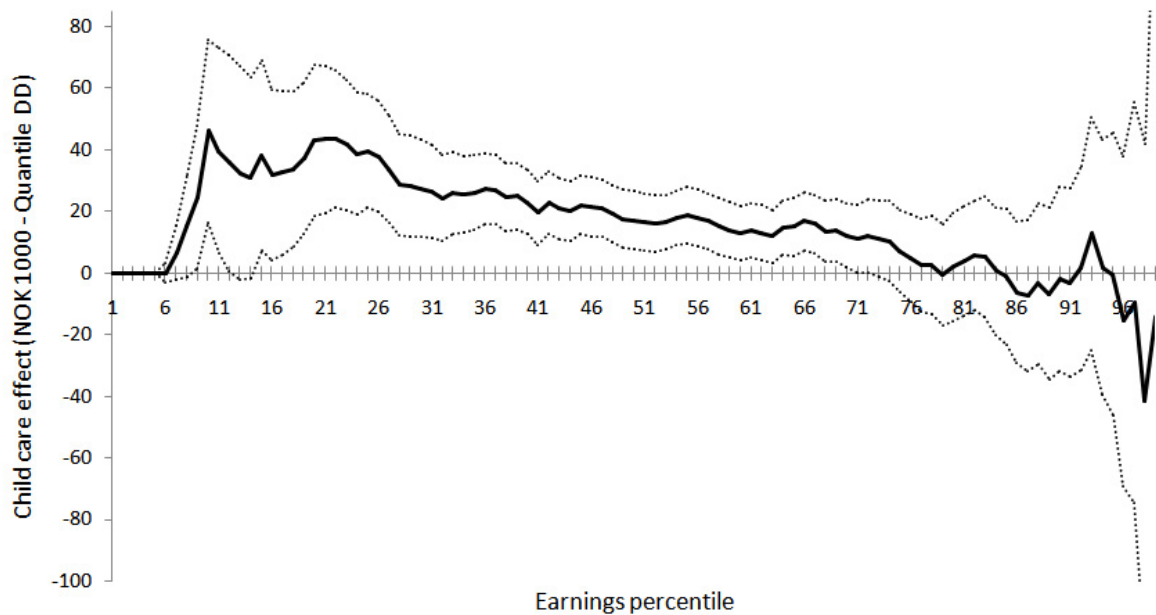
$$\tau^Q(q; X) = [\hat{F}_{11}^{-1}(q; X) - \hat{F}_{10}^{-1}(q; X)] - [\hat{F}_{01}^{-1}(q; X) - \hat{F}_{00}^{-1}(q; X)]. \quad (B2)$$

In our empirical analysis, we estimate the reform effects at all percentiles 1-99.¹ The standard errors are bootstrapped based on samples of $m = 10,000$ observations drawn with replacement from each of the pre- and post-reform cohorts. The standard deviations of the bootstrapped estimates are then adjusted by a factor of $\sqrt{2m/n}$, where n is the number of individuals in these cohorts (see Buchinsky,1995; Koenker, 2005).

¹Estimations are implemented using the *ivqte*-package detailed in Frölich, M. and B. Melly (2008). Estimation of quantile treatment effects with STATA. *Mimeo*.



(a) No covariates



(b) With covariates

Figure B1: Quantile DD estimates and 90%-confidence interval of the effect of the child care reform on the probability of earning above pre-treatment percentiles 1–99 without (panel a) and with (panel b) covariates.

Notes: Estimates are the difference between the estimated treatment effect from separate quantile regressions for pre- and post-reform cohorts, weighted by the propensity score, see Appendix B. Estimates and bootstrapped standard errors are corrected for intention to treat equal to .1785, see Section 4. The sample size is 498,956.