How Does Early Childcare Enrollment Affect Children, Parents, and Their Interactions?

By

Shintaro Yamaguchi, Yukiko Asai, Ryo Kambayashi

April 2018

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION
DISCUSSION PAPER NO. 29

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

How Does Early Childcare Enrollment Affect Children, Parents, and Their Interactions?*

Shintaro Yamaguchi† Yukiko Asai‡ Ryo Kambayashi§

August 10, 2017

Abstract

We estimate the effects of childcare enrollment on child outcomes by exploiting a staggered childcare expansion across regions in Japan. We find that childcare improves language development and reduces the symptoms of ADHD and aggression among the children of low-education mothers. Estimates show that the improved child behavior is strongly associated with better parenting quality, which seems to be brought about by informing mothers about good parenting practices and reducing parental stress. Our estimates for marginal treatment effects indicate that children who would benefit most from childcare are less likely to attend, implying inefficient allocation.

*The authors would like to thank Michihito Ando and seminar participants at the University of Toronto and the University of British Columbia for their helpful comments. All remaining errors are our own. Financial supports from the Murata Science Foundation and JSPS KAKENHI Grant Numbers 15K17071, 16H02020, 16K21743 are gratefully acknowledged. The data in the Longitudinal Survey of Newborns in the 21st Century were generously provided by the Japanese Ministry of Health, Labour and Welfare.

†Graduate School of Economics, the University of Tokyo. Hongo 7-3-1, Bunkyo, Tokyo, Japan. Email: syamaguchi@e.u-tokyo.ac.jp.

‡Institute of Social Science, the University of Tokyo. Hongo 7-3-1, Bunkyo, Tokyo, Japan. Email: y_asai@iss.u-tokyo.ac.jp.

§Institute of Economic Research, Hitotsubashi University. Naka 2-1, Kunitachi, Tokyo, Japan. Email: kambayas@ier.hit-u.ac.jp.

1 Introduction

Policymakers and experts consider early childhood education to be one of the most promising social programs. In 2002, the Barcelona European Council set objectives to improve the availability of high-quality and affordable childcare in the European Union. Similarly, in his 2013 State of the Union address, US President Obama called on Congress to extend access to high-quality preschool to all children.

Compelling evidence for the effectiveness of childcare is often from targeted programs such as the Perry Preschool Program in the US (see Heckman, Pinto, and Savelyev (2013)), but recent studies show that many universal childcare programs are also successful. For example, studies in Argentina (Berlinski, Galiani, and Gertler (2009)), Denmark (Datta Gupta and Simonsen (2010)), Germany (Cornelissen, Dustmann, Raute, and Schönberg (2015) and Felfe and Lalive (2015)), Norway (Havnes and Mogstad (2011, 2015) and Drange and Havnes (2015)), Spain (Felfe, Nollenberger, and Rodríguez-Planas (2015)), and the US (Cascio and Schanzenbach (2013)) all point to gains in child cognitive and/or non-cognitive outcomes from participation in early childhood education, and some identify stronger effects for disadvantaged children. However, little is known about how and why childcare enrollment improves child outcomes.

The objective of this paper is to estimate the effects of a large-scale nontargeted childcare program on children's cognitive and socioemotional skills, using Japanese data from the Longitudinal Survey of Newborns in the 21st Century (LSN21). We depart from previous work in this area by shedding light on the underlying mechanisms through which childcare influences children. We examine how childcare enrollment changes the inputs into children and the quality of inputs by estimating effects of childcare on the parenting quality, child-related expenses outside of childcare, parents' knowledge of good parenting practices, and parental stress and wellbeing, all of which may eventually also affect the children. Another contribution of this paper is that we provide new evidence on the efficacy of a childcare program on younger children than most previous studies. We analyze toddlers aged between 2 and 3, while most previous studies analyze preschool children aged between 4 and 5.

Childcare enrollment can influence children through two possible channels. The first channel is inputs into children including time at a childcare center, time with parents, and other market goods. Childcare enrollment increases time at a childcare center and likely to decrease their time with parents, which could increase or decrease child outcomes depending on the quality of time at home relative to that of time at childcare. Childcare centers in Japan are strictly regulated so that their quality is high and largely homogeneous across the country, and hence, quality of time at a childcare center may be better than that of time at home for disadvantaged children. If this is the case, childcare enrollment is likely to improve outcomes of disadvantaged children. However, childcare enrollment may not influence other children if they spend quality time at their home.

Childcare enrollment may also affect other child investment through market goods. Although it is heavily subsidized, childcare users pay about 28,400 JPY ($\approx$284 USD) per month, which may affect other expenses for children such as those on books. If parents perceive that time at a childcare center and other market goods are substitute for child development, the use of childcare crowds out inputs for children other than childcare. However, if parents perceive they are complement, childcare use can increase other inputs for children. The relationship between childcare and other child investment is not a priori known and an empirical question.

The second channel through which childcare affects children is the quality of parenting style. This channel could be interpreted as a technological change for the production function for a child's human capital. For a given set of inputs such as parental time with children, better quality parenting leads to better outcomes for children. Parenting quality is particularly relevant for children's socioemotional skill development. The studies from pediatrics and development psychology (see Gershoff (2002) and Deault (2010) for surveys) find that effective parenting practices can reduce children's behavioral problems.

There are at least two reasons for why childcare use can improve the quality of parenting style. First, parents may learn about better parenting practices from childcare teachers and other parents. Note that this effect is likely to vary by parents' educational background, because better educated mothers may learn about good parenting practices on their own, while other mothers do not. Second, childcare use may also reduce stress from raising a child so that mothers become patient and practice an effective parenting style. This may be the case given that mothers in Japan are the primary caregiver and often need to take care of her children all the time without much help from her spouse and other people. This effect is also likely to vary across mothers. If parenting skills are positively correlated with parents' education, those with low-education are likely to benefit more from childcare use, because they feel frustrated more than better-educated parents when not using childcare.

In the early 2000s, the national government initiated childcare reforms to increase female labor force participation and fertility rates. Although the national government covers half of the program cost, local governments are responsible for the rollout of the program. Depending on local governments' financial conditions and policy priorities, the pace of childcare expansion varies significantly. This has led to differences between regions across Japan, a feature we exploit for identification of the causal effects.

We estimate the treatment effects of childcare enrollment on various outcomes for children and their parents using instrumental variable (IV) regression, specifying childcare slots per child in a given region as an instrument. Because we control for year and region fixed effects in the estimation, the identifying variation is the regional variation in the rate of childcare expansion, which is similar to the difference-in-differences approach.

The estimates show that childcare enrollment has significant positive effects on language development and no effects on the symptoms of attention deficit hyperactivity disorder (ADHD) and aggression on average across children. However, when heterogeneous treatment effects are allowed for across mothers' educational background, we find that childcare enrollment reduces ADHD symptoms and aggression among children of low-education mothers relative to those of university-educated mothers. Children of low-education mothers demonstrate more ADHD symptoms and greater aggression if they are not enrolled in childcare, and hence, childcare enrollment reduces the gaps in behavioral outcomes among children.

To shed light on the mechanisms through which childcare influences children, we examine how parents change their behavior in response to childcare enrollment. The estimates show that childcare enrollment increases expenses for children besides childcare and improves low-education mother's parenting quality, both of which are expected to improve child outcomes. Further analysis suggests that a more positive home environment for children is brought about by informing low-education mothers about better parenting practices and reducing their stress from raising a child, although we cannot exclude the possibility that better behavior by children also improves their mothers' parenting practices. Given the strong and consistent association between child behavior and the home environment over a wide array of variables, our analysis suggests that a childcare program should educate not only children but also their parents, so as to promote their positive involvement.

Finally, we estimate the marginal treatment effect (MTE) that varies by the unobserved propensity for childcare use. The MTE framework (see Björklund and Moffitt (1987) and Heckman and Vytlacil (2005)) enables us to identify who would benefit most from childcare and how likely they are to use it. Our estimates indicate that while there are children who would particularly benefit from childcare enrollment, their mothers are less likely to use childcare. It is interesting to consider the characteristics of these mothers. Note that maternal employment is effectively a prerequisite for childcare access, and local governments will give higher priority to fulltime working mothers. Hence, the mothers who are less likely to use childcare tend to have weaker labor market attachment and lower labor market skills. Our counterfactual simulations indicate that while these mothers will eventually use childcare if its supply is sufficiently large, their responses to childcare reform are slow. Therefore, our analysis suggests that increasing the supply of childcare may not be enough, in itself, to improve childcare participation, and that other policy measures may also be necessary to bring the children of these mothers into formal childcare.

The remainder of the paper is structured as follows. Section 2 reviews the literature. Section 3 describes the institutional background and the childcare reform that we exploit for identification. Section 4 examines our household data. Section 5 outlines the econometric methods employed, including the MTE framework. Section 6 presents the estimation results for the IV regression and

Section 7 those for the MTE. Section 8 concludes.

2 Related Literature

This paper contributes to the literature on early childhood education. As mentioned in the introduction, there is a growing body of research on the efficacy of targeted and universal childcare programs across the world, and most of them examine children's test scores. Although the estimates are not directly comparable across studies, the test score effects appear to diminish with age. For example, in a US study, Cascio and Schanzenbach (2013) evaluate the effects of pre-K programs in Georgia and Oklahoma on test scores up until the eighth grade and find that any gains gradually decline.

Evidence for the effects on socioemotional skills is less extensive and the implications mixed. Baker, Gruber, and Milligan (2008) examine Quebec's universal childcare programs and find children are made worse in terms of aggression, motor and social skills, and health. Kottelenberg and Lehrer (2013) confirm that this negative finding is robust to the choice of statistical method and the cohort studied. Berlinski et al. (2009) estimate the effects of a universal program in Argentina and find positive effects on student self-control in the third grade, while Datta Gupta and Simonsen (2010) find that enrollment at 3 years of age in the universal childcare program in Denmark does not affect noncognitive outcomes at 7 years, regardless of the child's gender and her or his mother's education. Lastly, Felfe and Lalive (2015) examine the effect of childcare enrollment in Germany before 3 years of age on development at 6 years and conclude that enrollment improves the socioemotional development of children of less-educated mothers.

Evidence on the long-term effects of universal childcare programs is also scarce and mixed. Havnes and Mogstad (2011) find that childcare enrollment increases educational attainment and labor market participation and reduces welfare dependency in Norway. Estimates from this study, and later Havnes and Mogstad (2015), show that the Norwegian universal childcare program mostly benefits children from disadvantaged families, and does not have any positive impact on middle- and upper-class children. However, while their estimates are credible, they do not analyze how these positive outcomes are associated with the cognitive and socioemotional skills developed in formal childcare. Baker, Gruber, and Milligan (2015) consider the long-term consequences of the universal childcare program in Quebec, confirming their earlier finding of a negative impact on children's socioemotional skills. The results indicate that cohorts with increased childcare access subsequently have poorer health, lower life satisfaction, and higher crime rates later in life. Their analysis and evidence from targeted programs in the US (see Heckman et al. (2013)) show that it is the socioemotional skills that play a central role in long-term success in life.

While a number of studies estimate the efficacy of childcare programs, only a few examine the mechanisms through which childcare enrollment can impact children. Baker et al. (2008) find that the childcare program in Quebec leads to more hostile and less consistent parenting and lower-quality parental relationships. By contrast, Gelber and Isen (2013) analyze US data from the Head Start Impact Study and find that Head Start causes a substantial increase in parental involvement with their children. They find a positive association between children's cognitive test scores and parents' involvement with their children, however, the relationship between children's socioemotional skills and parents' time and other inputs into children is unknown. Herbst and Tekin (2010, 2014) analyze the effects of the Child Care and Development Fund, although this is not a childcare program but rather a childcare subsidy in the US. They find that receipt of the subsidy leads to lower child test scores, poorer child behavior, worse maternal health, and lower-quality interactions between parents and their children. Although the nature of the studied programs varies, all of these studies consistently indicate a strong association between child behavior, parenting quality, and maternal wellbeing. However, while the evidence suggests an important role for positive parenting practices, it is not entirely clear how childcare enrollment affects parenting quality.

There is a growing interest in heterogeneity in treatment effects. Bitler, Hoynes, and Domina (2014) estimate quantile treatment effects of Head Start on child outcomes and find that the gains are largest at the bottom of the skill distribution. Using data from Head Start Impact Study, Kline and Walters (2016) estimate the MTE of Head Start and its competing programs. They find that

Head Start generates largest test score for children who are less likely to participate in the program. Cornelissen et al. (2015) and Felfe and Lalive (2015) estimate the MTE of a childcare program in Germany. They find that while the average effects are insignificant, there are children who benefit from childcare enrollment. Cornelissen et al. (2015) find that many of these children are from disadvantaged and/or immigrant families and less likely to attend childcare. Our results echo these findings in that disadvantaged children benefit most from childcare, but they are less likely to attend childcare. We go a step further to uncover the underlying mechanisms by examining parental outcomes such as parenting quality.

3 Institutional Background

3.1 The Childcare System in Japan

Accredited Childcare Centers Some 94 percent of childcare centers in Japan satisfy the national quality standard set by the Child Welfare Act and are accredited by the governor of the province in which they are located[1]. In Japan, accredited childcare centers are subsidized by municipal, provincial, and national governments so that the average user pays only about 40% of the total cost. The average monthly fee per child is low at about 28,408 JPY ($\approx$ 284 USD), although this depends on the child's age, region, household income, and number of siblings. Because the vast majority of childcare centers are nationally accredited, and our main data set, LSN21, does not distinguish between accredited and nonaccredited centers, we refer to all nationally accredited childcare centers as childcare centers unless otherwise noted[2].

Quality The quality standard for childcare centers is established by the national government and is uniform across the country. In terms of the qualifications required by caregivers and the

[1]The actual administrative term used by the government is *prefecture*, but we use *province* as it is more intuitive for most readers.

[2]The remaining 6 percent are unaccredited childcare centers, which do not receive subsidies from the national government. However, many unaccredited childcare centers satisfy the quality standards set by local governments and receive subsidies accordingly.

child-to-caregiver ratio, Japan provides higher-quality childcare than most other countries in the OECD. Licensed caregivers have typically completed 2 years of postsecondary education, which is higher than elsewhere in the OECD (e.g. Germany, Norway, Sweden, and the Netherlands) where upper secondary education is required. There are three children per caregiver for children aged less than 1 year, and six for children aged between 1 and 2 years. These ratios are lower than in many comparable European countries. For example, the child-to-caregiver ratio in early childhood education is 16.2 in the UK, 12.0 in Denmark, 9.4 in Spain, 9.3 in Austria, 5.0 in Germany, and 4.8 in Sweden.

Program Intensity Most childcare programs are fulltime. Only 10 percent of enrolled children spend less than 7 hours per day in childcare, and most spend 7 to 10 hours. In addition, the vast majority attend childcare at least 5 days a week, with about 18 percent attending as much as 6 days a week. Only about 9 percent of children attend between 1 and 4 days a week. Detailed statistics are provided in online appendix.

Eligibility While the childcare program in Japan is not targeted at children from low-income households, neither is it quite universal. The Child Welfare Act imposes as an eligibility condition that parents and cohabiting adults must be unable to provide care for the child because of their usual work during the day, disability, sickness, pregnancy, participation in disaster restoration work, or other reasons approved by the local mayor. In practice, 94.2 percent of parents using a childcare center satisfy the eligibility condition on the basis of their usual work during the day.

Rationing Rule In Japan's major cities, notably Tokyo, the demand for subsidized childcare often exceeds supply. If this is the case, applications are ranked by need as assessed by the municipal government. Single parents are given highest priority and usually assigned a slot. Children from two-parent families are ranked highest when both parents work fulltime. A lower rank is given when at least one parent works less than fulltime. For example, the city of Yokohama assigns its highest rank (A) if both parents work at least 20 days per month and 40 hours per week, but its lowest rank (F) if at least one parent works for 16 days per month and 16–28 hours per week. Note that household income does not affect the rank, although it does change the fees to be paid.

3.2 Childcare Reform

The demand for subsidized childcare has exceeded supply in many regions because of the increase in the female labor supply since the early 1990s. Experts and policymakers believe that a lack of subsidized childcare increases the conflict between work and family life, and hence is responsible for Japan's low fertility rate.[3] The Basic Act for Measures to Cope with Society with Declining Birthrate was legislated in 2003, and the national government committed itself to taking legal and financial measures to increase the supply of childcare. As a result, between 2000 and 2010, the number of slots in childcare centers increased by about 12 percent, and the number of slots per child increased from 0.27 to 0.34. The rate of expansion in the number of childcare slots was slower than in other countries,[4] partly because the national government did not compromise its strict quality standards.

Note that the childcare slots are administratively determined and not estimated by the actual enrollment. Hence, we interpret the coverage rate as a measure of childcare supply relative to child population, because it does not pick up households' willingness to use.

3.2.1 Determinants of Rollout of the Reform

Even though the national government provides legal and financial support to expand the supply of childcare, the provincial and municipal governments are responsible for the rollout of the reform and need to match the funds made available by the national government. Because financial conditions and the policy priority placed on childcare vary between local governments, the pace of the rollout has differed considerably across regions. Specifically, there are three factors that can slow rollout of the childcare reform, according to the Cabinet Office (2010). First, the bureau-

[3] In 1990 the fertility rate was 1.54, and this had declined to 1.36 by 2000.

[4] For example, in Quebec, the total number of childcare slots more than doubled between 1997 and 2005. In Norway, the number of slots per child increased from 0.10 to 0.34 in the period 1975–1981.

cratic system may prevent local governments from acting in a timely manner. Second, some local governments do not have permanent funds to subsidize childcare centers. Third, suitable land and qualified childcare workers are scarce, particularly in major cities.

We assess how these factors and other regional characteristics affect the pace of childcare expansion by regressing the growth of the coverage rate from 2000 to 2010 on the pre-reform regional characteristics. The covariates include the female labor force participation rate, the total fertility rate, the financial capability index of the local government, the land price, and the average female wage in 2000.

We summarize the main regression result here and provide an extensive analysis in Appendix A. We find that the female labor force participation rate in the pre-reform period is positively correlated with the growth of the coverage rate, which suggests that the government increased the coverage rate in regions where the potential demand was high. The other factors do not have major effects on the growth of the coverage rate. This indicates that the growth of the coverage rate is not completely random, and hence, accounting for potential policy endogeneity is necessary for obtaining unbiased estimates. To address this issue, we include interactions between the pre-reform regional characteristics and the year dummies among our control variables. This modeling allows for flexible trends in outcome variables across regions so that the trends are correlated with the pre-reform regional characteristics.

3.2.2 Validity as Instrument

Our instrument is the childcare coverage rate, which is given by childcare slots per child in a region. Because the exogenous variables include year and region dummies as well as family characteristics, we account for time-constant differences across regions and nationwide changes in economic conditions and policies. Hence, this identification strategy is similar to the difference-in-differences approach.

In addition to accounting for potential policy endogeneity as discussed above, we address two more potential sources of endogeneity bias. First, the local economic conditions may affect parents' labor supply and hence, child outcomes. If they are correlated with the childcare coverage rate, our estimate would suffer an endogeneity bias. We avoid this problem by controlling for the province-level unemployment rates for population older than age 15.

Second, our estimates may be biased due to selective migration. Popular opinion is that obtaining a slot in an accredited childcare center is extremely difficult in Tokyo and that some parents even move to other districts just to access childcare. Using the Employment Status Survey 2012[5], we take a sample of mothers of children under 6 years old and examine the reasons for their most recent move, and where they moved from. We find that with respect to the reason "For childrearing and education", 9.5 percent moved within the same city, 4.6 percent moved from another city in the same province, and 1.4 percent moved from another province. Because we define a region in this study as a smaller geographic unit than province (see Section 4 below), this implies that at most 4.6 percent of the sample moved between regions for childcare purposes. As we show in Section 6.3, selective migration appears to have little effect on the estimates.

The exclusion restriction requires that the childcare coverage rate not affect the outcomes directly. It can be violated if the quality of childcare deteriorated during the childcare expansion. If the local governments traded off quality and quantity, the childcare expansion would then have had a negative effect on child development, which biases our estimates. However, this is unlikely, because the quality standard is legally set by the Child Welfare Act and is uniform across the country. Importantly, the national government did not change this regulation during the childcare reform.[6] Indeed, the regulation for quality control is partly responsible for the slow progress of the childcare reform relative to other countries.

Nevertheless, we address this issue by including a measure of childcare quality in the first and second stages of the IV regressions. Our measure of childcare quality is the number of children per teacher in accredited childcare centers in the region. Only teachers under a regular employment

[5] Conducted by the Statistics Bureau every 5 years and covering about 1 percent of the population.

[6] Local governments were originally not legally permitted to set a lower standard than the national standard. However, the Comprehensive Regional Sovereignty Reform was legislated in April 2011, which allowed local governments to set a lower standard. Nevertheless, the Ministry of Welfare, Labor and Health still publishes guidelines, and most local governments legislate their own standards following these guidelines.

contract are counted, and those under a temporary contract are not counted in the statistic. It is true that the child-teacher ratio measures only one of many aspects of childcare quality, but it is considered as one of the most relevant variables by experts. Indeed, many countries have regulations on the child-teacher ratio to control for childcare quality.

There are a few more issues that could affect our estimates. They include endogenous fertility, the presence of siblings, the choice of control variables, and the assumptions of the functional form. These issues are discussed extensively in Section 6.3, but we find our main results are largely unaffected.

4 Data

4.1 Data Sources

Our main data source is the LSN21, which is a census of children born in the periods January 10–17, 2001, July 10–17, 2001, and May 10–24, 2010. The first survey was conducted when the children were 6 months old and subsequent questionnaires were completed every year about 6 months after their birthdays. Surveys until the children were about 3½ years of age are currently available. The response rates were high at 93.5 and 88.1 percent in the first survey years for those cohorts born in 2001 and 2010, respectively. About 83 percent of respondents in the first survey remained in the survey at age 3½ years. These response rates are higher than those in the comparable National Longitudinal Survey of Children and Youth (NLSCY) conducted in Canada[7].

We draw data on accredited childcare centers from the annual Report on Social Welfare Administration and Services, which covers all provinces and major cities where the population exceeds 200,000 persons. We define a region as either a major city, or the set of all municipalities in a province except for the major cities. We include 82 regions covered in the data in both 2002 and 2011, which consist of 45 provinces and 37 major cities. The provinces of Fukushima and Miyagi

[7] In the NLSCY in Canada, the response rate in the first cycle conducted in 1994/95 was 86.5 percent, and 67.8 percent of children in the original cohort responded in the third cycle conducted in 1998/99.

are omitted owing to missing data because they were severely affected by the Great East Japan Earthquake and the ensuing tsunami in 2011.

The child population is from the quinquennial census. For years when the census is not available, we estimate the child population using linear interpolation. Other regional characteristics in 2000 are drawn from various sources. See the note accompanying Table 9 in Appendix A for details.

4.2 Variable Definitions

4.2.1 Treatment Variable

We define treatment by childcare enrollment at age 2½ years because child development outcomes are only available for children aged 2½ and 3½ years. While we do not control for enrollment status at other ages, the treatment and control groups exhibit very different childcare enrollment patterns over time as seen in Table 1.

At 6 months of age, few children are enrolled in childcare because many Japanese mothers are entitled to job-protected leave until their child reaches 1 year of age (see Asai (2015) and Yamaguchi (2016)). Many children begin attending childcare from the age of 1½ years. At this age, about 68 percent of treated children are enrolled, but only about 2 percent of the untreated. At age 3½ years, about 88 percent of treated children continue to attend childcare, while only about 14 percent of the untreated are enrolled. These statistics indicate that the enrollment patterns over time are very different for the treated and the untreated. Indeed at age 2½, the total years of childcare enrollment is 1.785 years for the treatment group, while it is only 0.027 years for the control group. The difference is even larger at age 3½.

4.2.2 Child Outcomes

We construct measures according to child language development, ADHD symptoms, and aggression. These measures are constructed from a set of questions that can be answered with a simple

Table 1: Childcare Enrollment Pattern Over Time

Age	0.5	1.5	2.5	3.5
Childcare Enrollment In a Given Year				
Treated	0.104	0.681	1.000	0.882
Control	0.004	0.023	0.000	0.135
Total Years of Childcare Enrollment				
Treated	0.104	0.785	1.785	2.668
Control	0.004	0.027	0.027	0.161

Source: LSN21.
Note: The treatment group is those who attend childcare at age 2½ and the control group is those who do not.

yes or no. Parents or other adults such as grandparents were eligible to respond, but in reality about 90% of the survey respondents were mothers.

We use the following three questions to measure the language development of 2½-year-old children: “Does your child say words such as ‘mom’?”, “Does your child put together two-word sentences?”, and “Does your child say his/her own name?” These are commonly used by pediatricians to measure child development and are included in the list of developmental milestones by the Centers for Disease Control and Prevention in the US.

We measure ADHD symptoms for 3½-year-old children using five questions comparable to those in the 5th edition of the Diagnostic and Statistical Manual (DSM), in the guidelines set by the American Psychiatric Association. The selected questions are: “Does your child listen until the other person has finished speaking?”, “Does your child cut in line?”, “Does your child scream in public spaces (e.g. buses, trains, and hospitals)?”, “Does your child have a short attention span?”, and “Is your child restless?” Similar measurements for ADHD are included in the NLSCY and analyzed by Baker et al. (2008) and Currie, Stabile, and Jones (2014). Aggression is another behavioral problem and is a part of disruptive behavior disorders, which closely resemble ADHD but are considered separate conditions by pediatricians. We measure child aggression at age 3½ years using the following three questions: “Does your child break books and toys?”, “Is your child violent?”, and “Is your child short-tempered?”

We construct indices for child outcomes by standardizing the number of positive responses in each category. These questions have some room for interpretation, and hence they are likely to be measured with error. We reduce the noise from measurement error by aggregating information from questions on the same theme. If responses are consistent over a set of questions, we consider the indices to reflect the child's actual behavior. We also verify that the indices are not driven by the response to a single question. As a result, all of our main results are relatively robust to the removal of any one variable from the set of variables that measure child development and behavior. Our main results are also unchanged if we use principal component analysis to construct the indices.

We normalize child outcome measures so that the mean is zero and the standard deviation is one. This normalization procedure is followed for outcomes for parents.

4.2.3 Outcomes for Parents

The index for parenting quality is constructed from responses to the question "How do you respond when your child behaves badly?" The five possible responses are: "Explain why your child should not do it", "Just say 'no' without explanation", "Ignore your child", "Spank your child", and "Confine your child in a place like a closet". For each of these, the respondent is asked to choose between "Always", "Sometimes", and "Never". These questions are asked when children are aged 3½ years.

We construct the parenting quality index by applying multiple correspondence analysis. This is a dimension reduction technique similar to principal component analysis and applicable to a set of ordered or categorical variables of the same substantive type. We summarize the main result and report the coordinates of each possible response in Table 11 in Appendix C. In the multiple correspondence analysis, the answers "Always explain", "Never say just 'no' without explanation", and "Never ignore the child" are regarded as indicators of high-quality parenting, while "Always confine the child in a place like a closet", "Always spank the child", and "Always ignore the child" are regarded as indicators of low-quality parenting.

Parental stress and subjective wellbeing is measured by the questions "What burdens do you carry when you raise your child?" and "What makes you happy when you raise your child?", respectively. The respondents are asked to select all that apply among 17 items for parental stress[8] and among 9 items for parental wellbeing.[9]

The survey also contains a few other variables relevant for determining the child's home environment. These ask whether parents know about good parenting practices, which is a self-reported binary response. The survey also requests the respondent to provide the childcare and nonchildcare expenses for the child in the survey month. Any expenses for the siblings of the child in question are excluded.

4.3 Descriptive Statistics

Table 2 presents summary statistics for the uptake, family characteristics, and outcomes variables. The enrollment rate for 2½-year-old children in childcare is 0.314. The mothers' labor market participation rate is 0.374, higher than the enrollment rate, suggesting that some mothers work using informal childcare arrangements.

The average age of mothers is 32.487 years, and fathers are about 2 years older. About 5 percent of mothers and fathers have less than a high school education, and about 35 percent graduated from high school without pursuing postsecondary education. Postsecondary education levels differ substantially between mothers and fathers. About 42 percent of mothers went to a 2-year college or equivalent, while only about 20 percent graduated from a 4-year university education or higher. By contrast, about 41 percent of fathers graduated from a 4-year university education and only 17 percent attended a 2-year college or equivalent.

There are slightly more boys than girls in the sample, and about 9 percent of the children had low birthweight as defined by the World Health Organization (less than 2,500 grams).

[8]They include "Fatigue", "Expenses for children", "Unable to have time for myself", "Spouse not cooperative", "Disagreement spouse about parenting", "Unable to work and/or do household chores", "Concerning about how others think about my child", "Need to keep an eye on child", "Not in a good relationship with parents of other children", etc.

[9]They include "Strengthened family ties", "Interactions with children", "Feeling that life is worthwhile", "Children interacting with each other", "Making more friends through raising the child", "Learning from the child", "The child making the whole family happier", "Growth of the child", and "Other".

Table 2: Summary Statistics

	All			Comparison by Treatment		
	Nobs.	Mean	S.D.	Treated	Untreated	p-value for Difference
Uptake Variables						
Childcare Enrollment	67913	0.314	0.464	1.000	0.000	
Market Work	67322	0.374	0.484	0.873	0.146	0.000
Hours of Work Per Week	67322	12.505	18.705	31.317	3.899	0.000
Mother's Characteristics						
Age	67913	32.487	4.482	32.612	32.430	0.000
Less Than High School	67913	0.042	0.201	0.043	0.042	0.778
High School	67913	0.331	0.470	0.304	0.343	0.000
2-Yr College	67913	0.423	0.494	0.416	0.426	0.016
4-Yr University or Higher	67913	0.204	0.403	0.236	0.189	0.000
Father's Characteristics						
Age	67913	34.482	5.485	34.508	34.470	0.404
Less Than High School	67913	0.069	0.254	0.084	0.062	0.000
High School	67913	0.350	0.477	0.365	0.343	0.000
2-Yr College	67913	0.169	0.375	0.183	0.163	0.000
4-Yr University or Higher	67913	0.412	0.492	0.369	0.432	0.000
Children's Characteristics						
Girl	67913	0.483	0.500	0.472	0.488	0.000
Low Birth Weight	67913	0.088	0.283	0.087	0.088	0.812
Child Outcomes						
Language Development	67510	0.000	1.000	0.153	−0.070	0.000
Aggression	61304	0.000	1.000	0.009	−0.004	0.152
ADHD Symptoms	59894	0.000	1.000	−0.035	0.016	0.000
Parent's Outcomes						
Parenting Quality	62140	0.000	1.000	0.010	−0.005	0.091
Lack of Parenting Knowledge	67728	0.089	0.284	0.084	0.091	0.002
Stress	67728	0.000	1.000	−0.034	0.016	0.000
Subjective Well-Being	67873	0.000	1.000	0.027	−0.012	0.000
Total Expenses	65999	2.602	2.853	4.431	1.751	0.000
Childcare Expenses	67547	0.904	1.676	2.708	0.089	0.000
Other Expenses	65740	1.677	2.198	1.714	1.660	0.003

Source: LSN21

Note: Children are in two-parent family. Child outcomes, parenting quality, and parent's stress and subjective well-being are normalized so that the mean is zero and standard deviation is one. Child aggression, ADHD symptoms, and parenting quality are measured when children are aged 3½ year old. Other variables are evaluated when children are 2½ year old. Expenses in the survey month are measured in 10,000 JPY (≈100 USD).

The total expenses for the surveyed children aged 2½ years are 26,020 JPY per month. This comprises monthly childcare expenses of 9,040 JPY, and other monthly expenses of 16,770 JPY. About 9 percent of mothers report that they do not know about good parenting practices.

We compare the characteristics and outcomes of treated and untreated families. The treatment in this study is enrollment at a childcare center at 2½ years of age. Given the large sample size, most of the differences are statistically significant, even if their magnitude is small.

Most treated mothers are in the labor market, as is expected, given the childcare eligibility rules. Treated mothers are better educated than untreated mothers, but treated fathers are less educated than untreated fathers. Skilled mothers have stronger labor market attachment, but the wives of skilled men are less likely to work. We do not find large differences in children's sex and birthweight by treatment status.

Treated children exhibit better language development and a lower frequency of ADHD symptoms than untreated children. We find no statistically significant difference in aggression. Treated parents report higher parenting quality, more parenting knowledge, lower parental stress, and better subjective wellbeing, although the differences are of small magnitudes. Expenses for surveyed children are greater for the treatment group. Most of the difference is due to childcare expenses; the difference in other expenses is small by treatment status.

Understanding the counterfactual care mode is important for interpreting the effects of center-based childcare, because the treatment effects are measured by the deviations from outcomes under that mode. Table 3 provides the distribution of childcare modes by the mothers' labor market status. The share of center-based childcare is 73 percent for working mothers. While this is the most common childcare mode for working mothers, many working mothers use other childcare modes. The next most common childcare mode is informal care by grandparents. Ten percent of working mothers do not report any nonparent childcare mode. The use of babysitters and other informal childcare is very rare, accounting for only 2 percent of all childcare.

The use of center-based childcare is uncommon for stay-at-home mothers. Its share of only 6 percent is reasonable because most parents need to satisfy the eligibility requirement by their usual work during the day. Most stay-at-home mothers do not report any nonparent childcare mode, and 15 percent use informal care by grandparents. The use of babysitters and other informal childcare is very rare among stay-at-home mothers.

Table 3: Childcare Mode by Mother's Labor Market Status

	Mother's Labor Market Status	
	Work	Home
Childcare Center	0.73	0.06
Grandparents	0.15	0.15
Sitters etc.	0.02	0.02
Parents Only	0.10	0.76

Source: LSN21

Note: All children are in two-parent family and 2½ years old. The primary childcare mode is mutually exclusive and collectively exhaustive as defined by the following rule. If enrollment for a childcare center is reported, this is considered as the primary mode, because most enrolled children attend full-time. If a child is cared by parents and grandparents only, the primary caregiver is grandparents. If any caregiver other than a childcare center and grandparents is reported, the primary caregiver is a child sitter. If no caregiver except for parents is reported, parents are the primary caregiver.

5 Econometric Methods

This section discusses our econometric methods. We first describe our specification for the IV regression, and then outline the MTE framework and the local IV estimator.

5.1 Instrumental Variable Regression

Our basic specification is based on IV regression. Define D as an indicator variable for childcare enrollment at age 2½ that takes a value of one if enrolled and zero if not. Let Y be an outcome variable and X be a K-dimensional vector of exogenous variables including year and region dummies and family and regional characteristics. The estimation equation is given by

$$Y = X\beta + \tau D + \varepsilon. \quad (1)$$

where ε is an error term that may be correlated with treatment status D. Assuming that the treatment effect is homogeneous, the parameter τ measures the treatment effect. When we estimate heterogeneous treatment effects varying by the mother's education, we interact the treatment status and dummy variables for the mother's education.

Childcare enrollment is determined by the following selection equation:

$$D = 1\{X\gamma + Z\delta > V\}, \quad (2)$$

where $1\{\cdot\}$ is an indicator function that takes a value of one if the condition in the curly brackets is satisfied and zero otherwise, Z is a vector of instrumental variables excluded from the outcome equation (1), and V is a scalar of unobserved characteristics. Our instrument Z includes the childcare coverage rate, which is defined as the number of childcare slots per child in a given region. We also include the interactions of the coverage rate and a subset of exogenous variables X in the instruments. We define the propensity score for childcare enrollment such that $P(X,Z) \equiv \Pr(D = 1|X,Z)$.

We estimate equation (1) using instruments 1, $P(X,Z)$, and X when we assume homogeneous treatment effects. We also use the interactions of $P(X,Z)$ and dummies for the mother's education as additional instruments when we allow for heterogeneous treatment effect varying by the mother's education. We use the propensity score $P(X,Z)$, instead of Z, as an instrument for the following reasons. First, this method produces the efficient IV estimator if the model for the propensity score is correctly specified. Second, this method is consistent, even if the propensity score is misspecified. Wooldridge (2010) extensively discusses these issues. Third, the IV estimate can be interpreted as a weighted average of the MTE with positive weights. We estimate the MTE by the local IV estimator using the propensity score to explore the role of unobserved heterogeneity. Our IV estimates can be interpreted in a unified framework when we use the propensity score as an instrument.

5.2 Marginal Treatment Effect

Define $j \in \{0,1\}$ as an index of enrollment status for childcare such that $j = 1$ indicates being enrolled at age 2½. A potential outcome Y_j for enrollment status j is given by

$$\begin{aligned} Y_j &= X\beta_j + U_j & (3) \\ E(U_j|X) &= 0, & (4) \end{aligned}$$

where U_j is an unobserved variable.

The enrollment status is determined by the selection equation (2) and can be rewritten as

$$\begin{aligned} D &= 1\{X\gamma + \delta Z > V\} & (5) \\ &= 1\{F_V(X\gamma + \delta Z) > F_V(V)\} & (6) \\ &= 1\{P(X\gamma + \delta Z) > U_D\}, & (7) \end{aligned}$$

where F_V is the cumulative distribution function for V, $P(\cdot)$ is the propensity score, and U_D is the quantile of unobserved characteristic V. We assume that (U_j, U_D) is independent of Z given X. We refer to U_D as the unobserved resistance to treatment, because a larger value of U_D keeps more families from treatment. This resistance to treatment summarizes all unobserved factors that determine the selection into treatment.

The MTE is defined as

$$MTE(X = x,\ U_D = u_D) \quad = \quad E(Y_1 - Y_0|X = x,\ U_D = u_D). \qquad (8)$$

This is interpreted as the gain from treatment for a family with observed characteristics $X = x$ and unobserved resistance to treatment $U_D = u$.

Heckman, Urzua, and Vytlacil (2006) show that the MTE can be estimated by the local IV estimator. We assume that the MTE is additively separable into an observed and an unobserved component,

$$MTE(X = x,\ U_D = u_D) \;=\; x(\beta_1 - \beta_0) + E(U_1 - U_0 | U_D = u_D). \tag{9}$$

The conditional mean outcome given the observed characteristics and the propensity score is

$$E(Y|X = x,\ P(X,Z) = p) \;=\; x\beta_0 + x(\beta_1 - \beta_0)p + K(p), \tag{10}$$

where $K(p)$ is a nonlinear function of the propensity score. The MTE for a family with $X = x$ and $U_D = p$ is given by the derivative of Equation (10) with respect to the propensity score,

$$\begin{aligned} MTE(X = x,\ U_D = p) &= \frac{\partial E(Y|X = x,\ P(X,Z) = p)}{\partial p} && (11)\\ &= x(\beta_1 - \beta_0) + \frac{\partial K(p)}{\partial p}. && (12) \end{aligned}$$

How does the local IV estimator identify the MTE defined by unobserved characteristics u_D? When the propensity score is p, those with the unobserved characteristics $u_D < p$ are selected into treatment and those with $u_D = p$ are indifferent. If we increase the propensity score by a small amount, those with $u_D = p$ are newly induced into the treatment. We can see the treatment effects on these newly treated persons by the change in the outcome in Equation (10) in response to the marginal change in p.

The unobserved resistance summarizes all the unobserved factors that determine childcare enrollment after observed characteristics such as parents' age and education are controlled. But what are they? Given the work requirement and the rationing rule that favors fulltime workers, the mother's labor market attachment and unobserved skills are likely to be the main components of

u_D. In the case of excess demand, the local government ranks applications by how much parents work. Because skilled mothers have a higher opportunity cost of staying at home, skilled mothers have a low u_D, so they are more likely to use childcare and work. By contrast, unskilled mothers are likely to have a high u_D.

6 Results

6.1 Childcare Enrollment

We identify the causal effects of childcare enrollment using regional variations in the childcare coverage rate. This identifying variation is graphically presented in Figure 1. We plot changes in the coverage rate over 2003–2012 on the horizontal axis and childcare enrollment for children aged 2½ years during the same period on the vertical axis for the 82 regions. The radii of the bubbles represent the number of observations.

As shown, the coverage rate increased in all regions during this period, but the magnitude of the changes varied considerably from 0.02 to 0.23. Childcare enrollment also increased in all regions, and the growth ranged from 0.01 to 0.27. The correlation coefficient is 0.62 and the standard error is 0.13, which is strongly significant. The graph provides prima facie evidence for the validity of our identification strategy.

We estimate the probability of childcare enrollment or the propensity score using the logit model. The covariates include the coverage rate up to the third-order polynomial, the parents' ages and education levels, the child's sex and birthweight, province-level unemployment rates, region-level child-teacher ratio, and dummies for year and region. The interactions between the coverage rate and parent characteristics are also included to allow for differential responses to the coverage rate. In addition, to address the possible policy endogeneity, we include interactions between the pre-reform regional characteristics in 2000 and the year dummies. All the parameter estimates except for the region fixed effects are in online appendix.

We focus on the selected variables and report their average marginal effects in Table 4. The

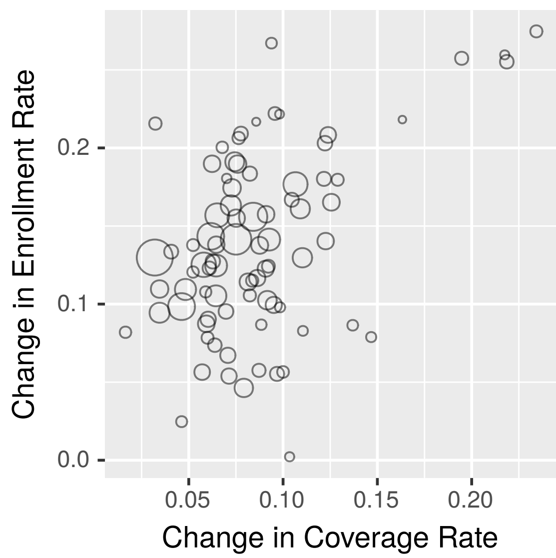

Figure 1: Growth of Coverage Rate and Enrollment Rate by Region

Source: LSN21, Census, and the annual Report on Social Welfare Administration and Services.
Note: Observations are for 2003-2012 and aggregated to the 82 regions. The radii of bubbles are the number of observations. The correlation coefficient is 0.62 and the standard error is 0.13.

average marginal effect of the coverage rate is significantly positive at 0.800. The high statistical significance gives us confidence in the validity of our IV regression.

The mother's age has a positive effect on childcare enrollment. The reference group for the mother's education is mothers graduating from a 4-year university education. The probability of childcare enrollment increases with the mother's education level, although the difference between high school graduates and those who did not graduate from high school is small. Overall, childcare enrollment increases with the mother's human capital. This is reasonable, because mothers are effectively required to be working as a precondition for the use of accredited childcare, and labor force attachment is stronger for skilled women.

The father's age has a significant negative effect on childcare enrollment. The probability of childcare enrollment decreases with the father's education level, although the difference between high school graduates and those with a 2-year college degree is small. The estimates indicate that childcare enrollment decreases with the father's human capital, which suggests that the father is the primary earner, and hence, his income has negative effects on both the mother's labor supply and childcare enrollment.

Table 4: First-Stage Regression: Average Marginal Effects on Childcare Enrollment at Age 2½

	Ave. Marginal Effect
Region	
Coverage Rate	0.800
	(0.136)
Mother	
Age	0.001
	(0.001)
Less Than High School	-0.098
	(0.010)
High School	-0.105
	(0.006)
2-Yr College	-0.072
	(0.005)
Father	
Age	-0.001
	(0.000)
Less Than High School	0.149
	(0.008)
High School	0.078
	(0.004)
2-Yr College	0.071
	(0.005)

Source: Authors' calculation from LSN21.
Note: The propensity score is estimated by the logit model. Standard errors are in parenthesis. The control variables include the coverage rate up to the third order polynomial, ages and education of parents, sex and birth weight of children, province-level unemployment rates, child-teacher ratio, and dummies for year and region. We also include interactions of pre-reform regional characteristics in 2000 and year dummies to account for the possible policy endogeneity. The coverage rate is interacted with characteristics of parents. Parameter estimates for the logit model are reported in online appendix.

6.2 IV Estimates

Using the IV estimator in Section 5.1, we provide the estimated treatment effects on several outcomes for both children and parents. The F-statistic for testing weak instrument is about 90, which implies that we can reject the null hypothesis that the instrument is weak. All the reported standard errors are clustered at the region level. For reader's convenience, we also report the results for reduced-form regressions in online appendix.

6.2.1 Maternal Labor Supply

Yamaguchi, Asai, and Kambayashi (2017) estimate the effects of childcare enrollment on various labor market outcomes for mothers. Because their specification is slightly different from that in this analysis, we re-estimate the models for the mother's labor market participation and weekly hours of work. The first two columns of Table 5 report the OLS and IV estimates for the mother's labor supply. The IV estimates without heterogeneity imply that childcare enrollment increases the mother's labor market participation by 65.5% and hours of work by 27.525 hours per week, which are similar to the OLS estimates.

We also estimate the heterogeneous treatment effects by the mother's education. The reference group here is mothers with 4-year university education or higher. The coefficient for childcare enrollment without interaction with other variables present the estimated effect for this reference group. Although some estimates are statistically significant, we hardly find a systematic pattern. Overall, the estimates show that childcare enrollment significantly increases the mother's labor market participation and hours of work regardless of her level of education. In the online appendix, we also report the estimated treatment effects by the mother's education, rather than the difference from the reference group.

Child Development and Behavior

We next provide estimates for child development and behavior from the third to fifth columns in Table 5. To facilitate interpretation, all of these outcome variables are normalized to have a mean of zero and a standard deviation of one.

The IV estimate without heterogeneity indicates that childcare enrollment increases the language development index by 0.595 standard deviations, which is statistically significant. We do not find evidence for treatment effect heterogeneity on language development, as shown by the coefficients for the interactions between childcare enrollment and the mother's education.

The evidence for the child's behavior from the IV estimate without heterogeneity is very weak.

Table 5: Effects of Childcare Enrollment on Mother's Work and Child Outcomes

	Mother's Labor Supply		Child Outcomes		
	Work	Hours	Language	ADHD	Aggression
OLS					
Childcare	0.723	27.172	0.227	-0.038	0.007
	(0.012)	(0.420)	(0.009)	(0.012)	(0.010)
IV (homogeneous effect)					
Childcare	0.655	27.525	0.595	-0.301	-0.054
	(0.083)	(3.164)	(0.191)	(0.265)	(0.210)
IV (heterogeneous effect)					
Less Than HS	-0.039	-1.825	-0.009	0.250	0.314
	(0.022)	(0.940)	(0.056)	(0.061)	(0.066)
High School	-0.079	-2.201	-0.004	0.090	0.088
	(0.017)	(0.688)	(0.042)	(0.056)	(0.042)
2-Yr College	-0.064	-1.636	0.039	0.039	0.050
	(0.013)	(0.556)	(0.034)	(0.049)	(0.044)
Childcare	0.573	29.111	0.686	-0.198	0.008
	(0.086)	(3.584)	(0.216)	(0.306)	(0.232)
Childcare $\times$ Less Than HS	-0.073	-6.053	-0.089	-0.378	-0.438
	(0.064)	(2.723)	(0.162)	(0.178)	(0.222)
Childcare $\times$ High School	0.081	-1.788	-0.073	-0.054	0.018
	(0.037)	(1.620)	(0.097)	(0.100)	(0.090)
Childcare $\times$ 2-Yr College	0.116	2.012	-0.107	-0.088	-0.102
	(0.030)	(1.515)	(0.091)	(0.092)	(0.101)

Note: Parameter estimates for Equation (1). Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity. The reference group for mother's education is mothers graduated from 4-year university or higher.

The IV estimate for ADHD symptoms is -0.301 and that for aggression is only -0.054. While the point estimates indicate that childcare enrollment reduce ADHD symptoms and aggression, neither of these estimates is statistically significant.

However, there is substantial heterogeneity in the treatment effects on the child's behavior. For children of low-education mothers, the treatment effects on ADHD symptoms and aggression are stronger than children of high-education mothers. The differences in the treatment effect on the ADHD and aggression indices are sizable and statistically significant at -0.378 and -0.438, respectively. The treatment effects for other children are not significantly different from the reference group.

These estimates imply that childcare enrollment reduces the gap in ADHD symptoms and aggression between children of low-education mothers and other children. When not enrolled in childcare, children of low-education mothers show much more ADHD symptoms and aggression than other children. Specifically, the indices of ADHD symptoms and aggression for children of low-education mothers are 0.250 and 0.314 standard deviations higher than the reference group, when not enrolled in children. However, these gaps completely disappear and the children of low-education mothers behave as well as the children of high-education mothers when enrolled in childcare.

We also report the treatment effects by the mother's education in the online appendix. The estimated effect on ADHD symptoms and aggression among children of low-education mothers are sizable and negative, but the latter is imprecisely estimated. Because the estimates for the difference in treatment effects are large and precise, our argument in the following focuses on the difference in treatment effects.

6.2.2 Parents' Outcomes

To understand the mechanism behind how childcare enrollment affects child outcomes, we examine the children's home environment and how it changes with childcare enrollment. We consider two channels here: parenting quality and monetary investment in children. The child development literature finds that parenting style affects children's behavior. In particular, pediatricians and development psychologists argue that corporal punishment leads to the child's problematic behavior. Parents can change their parenting styles for the following reasons. First, they may learn effective parenting styles from childcare teachers and possibly other parents using the same childcare center. This is particularly important for low-education mothers, because they would not learn about good parenting practices on their own, if not using childcare. Second, childcare use reduces stress from raising a child, which helps mothers avoid undesirable parenting style such as corporal punishment when her child does not behave well. This effect is also likely to be stronger for low-education mothers, because they would feel more stress from raising a child due to their low parenting skills.

Table 6: Effects of Childcare Enrollment on Parents' Outcomes

	Parenting Quality	Insufficient Parenting Knowledge	Stress	Subjective Well-Being	Childcare Expenses	Non-Childcare Expenses
OLS						
Childcare	-0.004	-0.009	-0.040	0.011	2.658	0.028
	(0.011)	(0.002)	(0.009)	(0.011)	(0.062)	(0.019)
IV (homogeneous effect)						
Childcare	0.111	0.022	-0.445	0.100	2.400	1.075
	(0.197)	(0.058)	(0.237)	(0.200)	(0.398)	(0.518)
IV (heterogeneous effect)						
Less Than HS	-0.510	0.043	0.153	-0.391	-0.161	-0.069
	(0.091)	(0.021)	(0.067)	(0.060)	(0.109)	(0.143)
High School	-0.300	0.033	0.020	-0.183	-0.181	-0.131
	(0.058)	(0.015)	(0.040)	(0.044)	(0.096)	(0.114)
2-Yr College	-0.154	0.031	-0.009	-0.021	-0.164	0.012
	(0.049)	(0.012)	(0.037)	(0.043)	(0.084)	(0.110)
Childcare	-0.115	0.099	-0.137	-0.007	2.980	0.812
	(0.273)	(0.074)	(0.258)	(0.235)	(0.492)	(0.597)
Childcare × Less Than HS	0.694	-0.150	-0.491	0.548	-0.740	0.658
	(0.245)	(0.055)	(0.185)	(0.171)	(0.286)	(0.384)
Childcare × High School	0.164	-0.059	-0.278	0.069	-0.570	0.273
	(0.118)	(0.032)	(0.083)	(0.093)	(0.181)	(0.256)
Childcare × 2-Yr College	0.104	-0.064	-0.152	-0.049	-0.237	-0.082
	(0.108)	(0.028)	(0.085)	(0.099)	(0.170)	(0.244)

Note: Parameter estimates for Equation (1). Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity. The reference group for mother's education is mothers graduated from 4-year university or higher.

Childcare may also affect children's home environment through household income. Although subsidized, the use of childcare is likely to increase the expenses for childcare. If childcare and other market goods are substitutes, childcare enrollment crowds out investment in children by reducing the purchase of market goods. If childcare and other market goods are complements instead, childcare enrollment increases investment in children by buying more market goods.

The first column in Table 6 shows how parenting quality changes by childcare enrollment. The IV estimates allowing for heterogeneous effects indicate that the effect of childcare enrollment is significantly stronger for low-education mothers. When not using childcare, parenting quality increases with the mother's education. The index for parenting quality is 0.510 standard deviations lower for low-education mothers than high-education mothers. Childcare enrollment does not affect high-education mother's parenting quality significantly, but the treatment effect for low-education mothers is 0.694 standard deviation greater. The estimates indicate that the use of childcare reduces the gaps in parenting quality among mothers with different educational background by improving low-education mother's parenting quality.

To examine potential reasons for why low-education mother's parenting quality improves, we estimate the effects of childcare enrollment on the mother's knowledge on parenting. The second column of Table 6 shows regression results in which the dependent variable is a binary indicator that takes one if the mother thinks she does not have sufficient parenting knowledge. Hence, a negative number in treatment effect means that childcare improves parenting knowledge. The estimate indicates that childcare use is more effective for low-education mothers than high-education mothers to improve their parenting knowledge.

The third and fourth columns presents estimates for the effects of childcare use on mother's stress and subjective well-being, which are self-reported. The estimates indicate that childcare use reduces stress and increases well-being of low-education mothers. Note that low-education mother's stress and well-being are worse than high-education mothers when not using childcare, but childcare use again eliminates the gaps between the two groups.

We next examine how childcare enrollment changes the family's expense on childcare and other expenses on the child. Note that these monetary expenses are specifically for the surveyed child, and any expenses for the child's siblings are not included. The IV estimate for the effect on childcare expenses without heterogeneity is positive and significant at 24,000 JPY ($\approx$240 USD). The treatment effect for low-education mothers is significantly lower, by 7,400 JPY, than that for high-education mothers, which is consistent with the fact that a lower fee is charged for low-income families.

The IV estimate for the effect on nonchildcare expenses without heterogeneity is positive and significant at 10,750 JPY ($\approx$107 USD), which implies that childcare and other child-related services and goods are gross complements, rather than substitutes. The treatment effect for low-education mothers is greater than for the reference group, although the estimate is imprecise.

We argue that parenting quality and nonchildcare expenses are likely to influence child outcomes, but that does not imply that the exclusion restriction is violated. The childcare reform (the IV) affected the probability of childcare enrollment, and childcare enrollment affects outcomes for both child and parents. We do not know exactly how child and parents' outcomes interact with each other, but we argue that the childcare reform does not directly affect any of these outcomes, but indirectly affects them through childcare enrollment.

6.3 Robustness

We now examine the robustness of our main results to alternative modeling assumptions. While there are many parameters in the model, we focus on just two key parameters: the estimate for treatment effect without heterogeneity and the difference in treatment effects between low- and high-education mothers. We address the following issues that may account for any biased estimates.

6.3.1 Potential Threats for Identification

Endogenous Fertility Our instrument is the coverage rate or the number of childcare slots per child, which is a measure of childcare availability. If childcare availability influences the fertility rate, the coverage rate is also affected. To avoid this potential problem from endogenous fertility, we estimate the model using an alternative measure of childcare availability, which is the number of childcare slots per woman aged 20–44 years in the region. Because it is plausible that the female population is more exogenous than the child population, the use of this alternative instrument helps us understand the extent of bias in the main specification.

Selective Migration Another threat to the identification is selective migration for childcare. Our estimates are upward biased if mothers who want to work move to a region where childcare is more readily available. Evidence from the Employment Structure Survey 2012 indicates that among mothers of children under 6 years of age, only 1.4 percent move from another province "For childrearing and education". Because interprovincial migration for childrearing and education is very uncommon, the estimates based on the province-level variables are unlikely to be biased by selective migration. We assess the extent of possible bias due to selective migration by comparing our preferred estimates with those of an alternative model in which the region-level variables are aggregated at the province level.

Siblings In our preferred specification, we do not control for the number of siblings, because it may be affected by the availability of childcare. However, child and parental outcomes may vary by the number of children in the household. In addition, the municipal government prioritizes a family for childcare if an older sibling is already enrolled in the same childcare center. Omitting the number of siblings therefore does not bias the results, but including it may lead to precise estimates as long as it is exogenous. To address this, we augment the main specification by adding the numbers of younger and older siblings to the first- and second-stage regressions.

Coverage Rate in Earlier Years Our treatment variable is an indicator for childcare enrollment at age 2½ years and instrumented by the coverage rate at that age. However, 68 percent of treated children were already enrolled in childcare 1 year before. This implies that the coverage rates in previous years may also be relevant for predicting childcare enrollment. Leaving out the coverage rates in earlier years does not bias our estimates, but including them in the first-stage regression increases the efficiency of the estimator. We address this by estimating the propensity score using the logit model that includes up to the third-order polynomials of the coverage rates at ½ and 1½ years. All covariates in the benchmark specification are included, and the specification for the second-stage regression remains identical.

Linear Probability Model in the First-Stage Regression Our preferred model for the propensity score is the logit model because it ensures the predicted scores lie between zero and one. The propensity score must be correctly specified for consistent estimation of the MTE by the local IV estimator. However, the propensity score does not have to be correctly specified for the IV regression. Because the logit model is a nonlinear estimator, our estimates may be driven by nonlinearity, rather than variation in the data. To assess the consequences of the use of a nonlinear model, we estimate the propensity score using the linear probability model and specify it as an instrument in the second-stage regression.

6.3.2 Results

The estimates for the alternative models are reported in Table 7. The estimates for the baseline specification are also reproduced for convenience. Most estimates for the alternative models are very similar to those of the baseline model, but there are some noticeable differences. When the variables are aggregated at the province level to address selective migration (Model 3), the IV estimate for the treatment effect on nonchildcare expenses is small and insignificant. When the number of siblings is included (Model 4), the estimated treatment effects on the mother's participation and hours of work are implausibly greater than the benchmark estimates. Overall, our estimates are robust to the alternative modeling assumptions.

Table 7: Robustness Checks

	(1) Benchmark	(2) Endogenous Fertility	(3) Selective Migration	(4) Siblings	(5) Lagged Cov. Rate	(6) Linear in 1st Stage
Homogeneous Effects						
Participation	0.655	0.671	0.796	1.077	0.646	0.662
	(0.083)	(0.091)	(0.053)	(0.093)	(0.081)	(0.095)
Hours	27.525	26.083	27.083	44.457	25.897	24.763
	(3.164)	(3.204)	(1.828)	(3.740)	(3.575)	(3.587)
Language Development	0.595	0.553	0.348	0.585	0.458	0.497
	(0.191)	(0.198)	(0.105)	(0.163)	(0.182)	(0.202)
Aggression	-0.301	-0.434	0.043	-0.221	-0.315	-0.111
	(0.265)	(0.267)	(0.123)	(0.241)	(0.266)	(0.276)
ADHD Symptoms	-0.054	-0.146	0.101	-0.100	-0.113	0.074
	(0.210)	(0.205)	(0.131)	(0.217)	(0.207)	(0.233)
Parenting Quality	0.111	0.182	-0.108	-0.102	0.077	0.271
	(0.197)	(0.196)	(0.138)	(0.164)	(0.187)	(0.241)
Lacking Parenting Knowledge	0.022	0.034	0.003	-0.002	0.016	-0.017
	(0.058)	(0.060)	(0.034)	(0.064)	(0.056)	(0.059)
Stress	-0.445	-0.333	-0.221	-0.491	-0.433	-0.424
	(0.237)	(0.238)	(0.133)	(0.237)	(0.223)	(0.250)
Well-Being	0.100	-0.105	0.213	0.003	0.075	0.196
	(0.200)	(0.203)	(0.127)	(0.196)	(0.195)	(0.214)
Childcare Expense	2.400	3.015	1.997	1.698	2.479	2.461
	(0.398)	(0.452)	(0.269)	(0.318)	(0.387)	(0.385)
Other Expense	1.075	1.325	0.170	0.655	1.043	0.981
	(0.518)	(0.565)	(0.285)	(0.381)	(0.509)	(0.487)
Diff. in Effects for Low-Educ. Mothers						
Participation	-0.073	-0.076	-0.076	-0.132	-0.076	-0.075
	(0.064)	(0.064)	(0.066)	(0.064)	(0.064)	(0.065)
Hours	-6.053	-5.841	-6.254	-9.366	-6.133	-5.970
	(2.723)	(2.685)	(2.892)	(2.720)	(2.692)	(2.727)
Language Development	-0.089	-0.089	-0.073	0.004	-0.070	-0.088
	(0.162)	(0.161)	(0.167)	(0.138)	(0.158)	(0.161)
Aggression	-0.378	-0.385	-0.428	-0.318	-0.372	-0.371
	(0.178)	(0.182)	(0.182)	(0.166)	(0.177)	(0.177)
ADHD Symptoms	-0.438	-0.435	-0.461	-0.383	-0.436	-0.411
	(0.222)	(0.224)	(0.213)	(0.204)	(0.218)	(0.226)
Parenting Quality	0.694	0.705	0.661	0.636	0.695	0.711
	(0.245)	(0.247)	(0.252)	(0.211)	(0.244)	(0.247)
Lacking Parenting Knowledge	-0.150	-0.152	-0.147	-0.102	-0.148	-0.150
	(0.055)	(0.054)	(0.050)	(0.051)	(0.054)	(0.056)
Stress	-0.491	-0.515	-0.482	-0.356	-0.478	-0.486
	(0.185)	(0.186)	(0.173)	(0.182)	(0.186)	(0.183)
Well-Being	0.548	0.576	0.537	0.520	0.546	0.570
	(0.171)	(0.165)	(0.161)	(0.160)	(0.168)	(0.176)
Childcare Expense	-0.740	-0.813	-0.715	-0.612	-0.733	-0.739
	(0.286)	(0.322)	(0.271)	(0.238)	(0.289)	(0.307)
Other Expense	0.658	0.631	0.617	0.771	0.654	0.671
	(0.384)	(0.385)	(0.417)	(0.306)	(0.382)	(0.376)

Note: See note for Table 5 for the specification of the benchmark model. Standard errors are in parenthesis and clustered at the region level. In model (2), the instrument is the number of childcare spots per women aged 20-44. In model (3), regional variables are aggregated to the province level. In model (4), the numbers of younger and older siblings are included. In model (5), the coverage rates at age ½ and 1½ are also included in the set of instruments. In model (6), the linear probability model instead of the logit model is used in the first stage.

6.4 Interpretation and Discussion

6.4.1 Child Outcomes

The estimates indicate that childcare enrollment on average improves the language development of children and the treatment effects are homogeneous across children with different mother's educational background. Note that the language development index is constructed from questions for which the response would be positive for most children, e.g. 88 percent of children can say words such as "mom", put together two-word sentences, and say their own names. This suggests that our index is useful for detecting children with substantially slower language development, rather than ranking children. Hence, even if childcare enrollment improves the language development of above-average children, we are unable to detect such effects owing to the nature of the index. In this sense, our index is similar to the German school readiness test, which 91 percent of German children pass (see Cornelissen et al. (2015)).

Our estimates indicate that the effects on child aggression and ADHD symptoms are almost zero on average. However, childcare enrollment significantly improves the behavior of children of low-education mothers. They behave worse than those of high-education mothers when not enrolled in childcare, but they behave equally well when enrolled in childcare. This result is robust to alternative modeling assumptions.

The indices for children's behavior are less objective than the language development index, and we cannot rule out the possibility that outcomes are measured with error. We attempt to minimize the role of measurement errors by using aggregated indices, rather than relying on a single variable. Note that this measurement error issue is not unique to our study; it is also found in previous work by Baker et al. (2008) and Currie et al. (2014) using the NLSCY.

6.4.2 Understanding Mechanisms

Childcare enrollment can affects children's cognitive and socioemotional skills by changing inputs into children. It is plausible that the learning environment at a childcare center may be of better quality than the home environment of some children. If so, spending more time at a childcare center can directly improve these outcomes. In addition, childcare enrollment may improve child outcomes indirectly through encouraging a better home environment. Our estimates indicate that childcare enrollment improves parenting quality, nonchildcare expenses, parenting knowledge, and the wellbeing of low-education mothers, which may eventually affect their children.

Although few economics studies explicitly consider this indirect path, there is theory and empirical evidence from pediatrics and developmental psychology of the effects of parenting practices on child behavioral outcomes (see Gershoff (2002) and Deault (2010) for surveys). While many studies adopt a correlational or longitudinal design, Shaw et al. (2006), Gardner et al. (2007), and Brotman et al. (2011) establish a causal effect from randomized controlled trials that are designed to promote effective parenting practices. These studies find that the interventions improve parenting quality and the behavior of children from disadvantaged families.

Although our results are consistent with these findings from the field of developmental psychology, we do not rule out the possibility that better child development and behavior also improve parenting quality. Indeed, Deault (2010) reviews empirical studies on the association between ADHD and parenting practices and finds that poor child behavior can lead to poor parent–child interaction and harm to the mother's mental health. Nevertheless, our estimates indicate that childcare enrollment improves the parenting knowledge of low-education mothers, which is unlikely to be driven by any change in child behavior. Overall, our analysis suggests that educating not only children but also their parents can increase the effectiveness of a childcare program.

7 Marginal Treatment Effects

7.1 Local Instrumental Variable Estimates

We estimate the MTE, which varies by unobserved resistance to treatment. Given the work requirement and the rationing rule that favors fulltime workers, the unobserved resistance is likely to represent the mother's unobserved labor market attachment and skills after their age and education are controlled. Namely, those with weak resistance to treatment are likely to be skilled, while those with strong resistance to treatment are likely to be unskilled. The MTE estimates enable us to understand how treatment effects vary by such unobserved characteristics.

Figure 2 depicts how the MTE on child development and behavioral outcomes changes with unobserved resistance to treatment u_D. Note that negative values for treatment effects on aggression and ADHD symptoms imply that treatment improves children's behavior (i.e. less bad behavior). For all of the three child outcomes, children from families with a weak resistance to treatment are not significantly affected by childcare enrollment. By contrast, childcare enrollment improves the outcomes of children from families with a strong resistance to treatment. These results are in line with the finding about children of low-education mothers, because mothers with a strong resistance to treatment are likely to be unskilled.

The MTE curves for parental outcomes are reported in online appendix. They tend to be noisy and statistically insignificant. We cannot observe a systematic relationship between unobserved resistance and these parental outcomes.

We also calculate the ATE, treatment effect on the treated (TT), and treatment effect on the untreated (TUT) by taking the relevant weighted averages of the MTE. The estimates for the aggregate treatment parameters and the corresponding weights are shown in online appendix. The ATEs are similar to those estimated by the IV regression. For outcomes for which the MTE curve is upward sloping, the TT is smaller than the TUT, and vice versa when the MTE curve is downward sloping.

7.2 Counterfactual Policy Simulations

To evaluate the effects of a further childcare expansion, we conduct counterfactual simulations in which the coverage rate is raised by 0.1, 0.2, or 0.3. Table 8 summarizes the simulation results. Given the current coverage rate, the childcare enrollment rate is 0.392. This increases to 0.480 when the coverage rate is raised by 0.1. When the coverage rate is raised by 0.3, the enrollment rate increases to 0.649, which is about the same as the enrollment rate for children aged 0–2 years

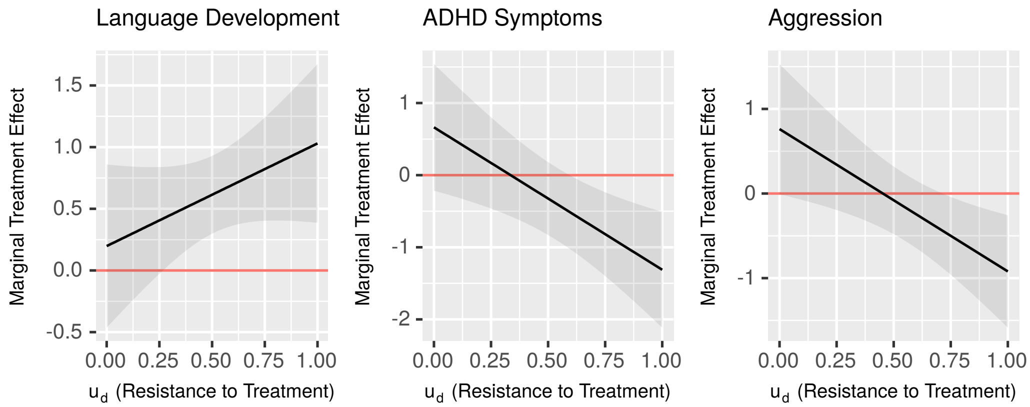

Figure 2: Marginal Treatment Effect Curves

Note: The marginal treatment effect is graphically presented with the 90% confidence interval. The standard errors are clustered at the region level. The MTE curve is based on the estimated outcome equation (10) and evaluated at the mean of all covariates except for the propensity score.

in formal childcare in Denmark, which has the highest childcare enrollment rate in the OECD.

Table 8: Counterfactual Policy Simulations

	Policy-Relevant Treatment Effect			**Propensity Score**	
	Mental Development	Aggression	ADHD	Baseline	New Policy
Raise Coverage Rate by 0.1	0.559	-0.179	0.040	0.392	0.481
	(0.182)	(0.296)	(0.218)	(0.012)	(0.020)
Raise Coverage Rate by 0.2	0.593	-0.259	-0.029	0.392	0.568
	(0.175)	(0.289)	(0.210)	(0.012)	(0.032)
Raise Coverage Rate by 0.3	0.625	-0.334	-0.093	0.392	0.648
	(0.171)	(0.287)	(0.207)	(0.012)	(0.044)

Note: Simulations are based on the estimated MTE. Standard errors are in parenthesis and clustered at the region-level. Weights for the policy-relevant treatment effect are provided in online appendix.

We calculate the policy-relevant treatment effects on three child outcomes. The policy-relevant treatment effects are average effects for those newly induced into treatment by the policy change. As shown in Appendix B, the policy-relevant treatment effects are given by the weighted average of the MTE. The estimates indicate that as the coverage rate increases, the policy-relevant treatment effects become increasingly strong in the direction that improves child outcomes. The MTE curves for child outcomes show that the MTE is stronger for children of parents with a stronger resistance to treatment, that is, those who are less likely to use childcare. As the childcare coverage rate increases, these children are gradually enrolled in childcare and improve their behavior.

The MTE estimates indicate that children who would most benefit from childcare enrollment are less likely to be enrolled. These children are gradually induced into treatment as childcare reforms progress; however, their enrollment levels are slow to respond to the expanded supply. Our analysis therefore suggests that increasing the supply of childcare may not produce sufficiently fast results among the population segment that would benefit most from formal childcare, and that other policy measures are also necessary to bring these children into formal childcare. The efficacy of the program could be improved by targeting children from disadvantaged families.

8 Conclusion

We estimate the effects of childcare enrollment on the outcomes of children and parents. Our estimates indicate that childcare enrollment improves language development and reduces aggression and ADHD symptoms in children of low-education mothers. These children show more ADHD symptoms and aggression than other children when they are not enrolled in childcare. However, childcare enrollment helps them catch up and provides a level playing field.

Part of the strong positive effect for children of low-education mothers may be brought about through better parenting quality and greater monetary investment in children. Although we cannot exclude the possible effects of child behavior on parenting quality, evidence suggests that childcare enrollment informs low-education mothers about good parenting practices and reduces their stress from raising a child, which may in turn improve their parenting quality. Promoting positive parental involvement therefore might further improve the effectiveness of a childcare program.

The MTE framework enables us to identify the treatment effects varying by the unobserved propensity to use childcare. The estimates indicate that childcare enrollment is effective for some children, but their mothers are less likely to use childcare. Because the rationing rule ranks childcare applications by how much the parents work, the mothers of nonparticipants are more likely to have weak labor force attachment and low skills. This implies that the rationing rule may prevent disadvantaged children from being enrolled in childcare. Although the rationing rule is different in other countries, tax deductions for childcare are commonly available in most. Such deductions are also likely to lead to negative selection into treatment, similar to the pattern found in this analysis, because they lower the effective price of childcare only when both parents work and pay significant tax. Our analysis suggests that childcare and other related social programs need to be carefully designed to ensure that this public service is delivered to children from disadvantaged families.

In terms of limitations, our measures of child outcomes and parenting quality are based on simple yes/no answers to questions that may leave room for interpretation in some cases, so we are unable to exclude the possibility that outcomes are measured with significant errors. Another limitation is that our outcome measures are contemporaneous with or 1 year after childcare enrollment, and the long-term outcomes of childcare enrollment remain largely unknown. The positive effects of childcare may either dissipate over time or persist into adulthood, as shown by existing studies. These important issues are left for future research.

References

ASAI, Y. (2015): "Parental Leave Reforms And The Employment of New Mothers: Quasi-Experimental Evidence from Japan," *Labour Economics*, 36, 72–83.

BAKER, M., J. GRUBER, AND K. MILLIGAN (2008): "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 116, 709–745.

——— (2015): "Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program," Working Paper 21571, National Bureau of Economic Research.

BERLINSKI, S., S. GALIANI, AND P. GERTLER (2009): "The Effect of Pre-Primary Education on Primary School Performance," *Journal of Public Economics*, 93, 219–234.

BITLER, M. P., H. W. HOYNES, AND T. DOMINA (2014): "Experimental Evidence on Distributional Effects of Head Start," Working Paper 20434, National Bureau of Economic Research.

BJÖRKLUND, A. AND R. MOFFITT (1987): "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *The Review of Economics and Statistics*, 69, 42–49.

BROTMAN, L. M. ET AL. (2011): "Promoting Effective Parenting Practices and Preventing Child Behavior Problems in School Among Ethnically Diverse Families From Underserved, Urban Communities," *Child Development*, 82, 258–276.

CABINET OFFICE (2010): "A Project to Reduce Wait-Listed Children Through Cooperation Between the National and Local Governments," Available from http://www.kantei.go.jp/jp/singi/taikijidou/kettei/kihon.pdf.

CASCIO, E. AND D. W. SCHANZENBACH (2013): "The Impacts of Expanding Access to High-Quality Preschool Education," *Brookings Papers on Economic Activity*, 47, 127–192.

CORNELISSEN, T., C. DUSTMANN, A. RAUTE, AND U. SCHÖNBERG (2015): "Who benefits from universal childcare? Estimating marginal returns to early chilcare attendance," University College London.

——— (2016): "From LATE to MTE: Alternative Methods for The Evaluation of Policy Interventions," *Labour Economics*, 41, 47 – 60, sOLE/EALE Conference Issue 2015.

CURRIE, J., M. STABILE, AND L. JONES (2014): "Do Stimulant Medications Improve Educational and Behavioral Outcomes for Children With ADHD?" *Journal of Health Economics*, 37, 58–69.

DATTA GUPTA, N. AND M. SIMONSEN (2010): "Non-cognitive child outcomes and universal high quality child care," *Journal of Public Economics*, 94, 30–43.

DEAULT, L. C. (2010): "A Systematic Review of Parenting in Relation to The Development of Comorbidities and Functional Impairments in Children With Attention-Deficit/Hyperactivity Disorder (ADHD)," *Child Psychiatry & Human Development*, 41, 168–192.

DRANGE, N. AND T. HAVNES (2015): "Child Care Before Age Two and the Development of Language and Numeracy: Evidence from a Lottery," IZA Discussion Paper 8904.

FELFE, C. AND R. LALIVE (2015): "Does Early Child Care Affect Children's Development," University of Lausanne.

FELFE, C., N. NOLLENBERGER, AND N. RODRÍGUEZ-PLANAS (2015): "Can't Buy Mommy's Love? Universal Childcare and Children's Long-term Cognitive Development," *Journal of Population Economics*, 28, 393–422.

GARDNER, F. ET AL. (2007): "Randomized Prevention Trial for Early Conduct Problems: Effects on Proactive Parenting and Links to Toddler Disruptive Behavior." *Journal of Family Psychology*, 21, 398–406.

GELBER, A. AND A. ISEN (2013): "Children's Schooling and Parents' Behavior: Evidence from the Head Start Impact Study," *Journal of Public Economics*, 101, 25–38.

GERSHOFF, E. T. (2002): "Corporal Punishment by Parents and Associated Child Behaviors and Experiences: a Meta-Analytic and Theoretical Review." *Psychological Bulletin*, 128, 539–579.

HAVNES, T. AND M. MOGSTAD (2011): "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes," *American Economic Journal: Economic Policy*, 3, 97–129.

——— (2015): "Is universal child care leveling the playing field?" *Journal of Public Economics*, 127, 100 – 114, the Nordic Model.

HECKMAN, J., R. PINTO, AND P. SAVELYEV (2013): "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *The American Economic Review*, 103, 2052–2086.

HECKMAN, J. J., S. URZUA, AND E. VYTLACIL (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 88, 389–432.

HECKMAN, J. J. AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669–738.

HERBST, C. M. AND E. TEKIN (2010): "Child Care Subsidies and Child Development," *Economics of Education Review*, 29, 618–638.

——— (2014): "Child Care Subsidies, Maternal Health, and Child–Parent Interactions: Evidence from Three Nationally Representative Datasets," *Health Economics*, 23, 894–916.

KLINE, P. AND C. R. WALTERS (2016): "Evaluating Public Programs with Close Substitutes: The Case of Head Start," *The Quarterly Journal of Economics*, 131, 1795–1848.

KOTTELENBERG, M. J. AND S. F. LEHRER (2013): "New Evidence on the Impacts of Access to and Attending Universal Child-Care in Canada," *Canadian Public Policy*, 39, 263–286.

SHAW, D. S. ET AL. (2006): "Randomized Trial of a Family-Centered Approach to The Prevention of Early Conduct Problems: 2-Year Effects of The Family Check-Up in Early Childhood." *Journal of consulting and clinical psychology*, 74, 1–9.

WOOLDRIDGE, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*, vol. 1 of *MIT Press Books*, The MIT Press.

YAMAGUCHI, S. (2016): "Effects of Parental Leave Policies on Female Career and Fertility Choices," McMaster University.

YAMAGUCHI, S., Y. ASAI, AND R. KAMBAYASHI (2017): "Effects of Subsidized Childcare on Mother's Employment Under a Rationing Mechanism," McMaster University.

A Determinants of the Growth the Supply of Childcare (For Online Publication)

Cabinet Office (2010) argues that three factors slow down the rollout of the childcare reform. First, the bureaucratic system prevents them from acting timely. Second, some local governments do not have permanent funds to subsidize childcare centers. Third, land and qualified childcare workers are scarce particularly in major cities.

We assess how these factors and other regional characteristics affect the pace of childcare expansion. Following the literature, we measure the supply of childcare relative to potential demand by the coverage rate defined by the number of spots per children aged 0-5 in a given region. By construction, the coverage rate increases when the number of spots increases, the number of children decreases, or both happen. Let C_t and N_t be the numbers of spots and children, respectively. The change of the coverage rate from $t+1$ to t can be decomposed as

$$\frac{C_{t+1}}{N_{t+1}} - \frac{C_t}{N_t} = \left[\frac{C_{t+1} - C_t}{N_t}\right] - \left[\frac{C_{t+1}}{N_t} - \frac{C_{t+1}}{N_{t+1}}\right]. \tag{13}$$

We refer the first term on the right hand side as the supply factor and the second term as the population factor. The first term measures the growth of the number of spots per children in the base year t, and hence, this term isolates the effect of a change in the supply of childcare. The second term measures the effect of a change in child population. With a fixed number of childcare spots, fewer children implies a higher coverage rate.

We regress the supply and population factors as well as the change of the coverage rate from 2000 to 2010 on regional characteristics for 82 regions used in our main analysis.[10] Given the objectives of the childcare reforms and the argument above, we include the female labor force participation rate, the total fertility rate, the financial capability index of the local government, land price, and average female wage in 2000. These factors are likely to influence the decisions on childcare supply. Although these factors may not causally affect the population size of children,

[10] See Section 4 for our definition of region and selection criteria.

they may be correlated. Note that the population size of children is indirectly affected by the population size of young adults. Young adults tend to move from rural regions or smaller cities to major cities for school and/or work,[11] which may eventually affect the size of child population.

Table 9 reports the regression results. The first column shows determinants of the supply of childcare. The female labor force participation rate in 2000 is negatively correlated with the growth of the supply of childcare. This implies that the supply of childcare increased in the regions where the female labor force participation rate was low, which is consistent with the objective of the childcare reform. The effect of total fertility rate is small and insignificant. As expected, the effect of financial capability of local governments is positive, while the effects of land prices and wages of female workers are negative but insignificant.

Table 9: Determinants of the Growth of Childcare Coverage Rate in 2000-2010

	Change in Supply	Change in Child Population	Change in Coverage Rate
Female Labor Force Participation Rate	−0.557	−0.743	0.186
	(0.190)	(0.156)	(0.089)
Total Fertility Rate	0.154	0.156	−0.001
	(0.106)	(0.087)	(0.050)
Financial Capability Index	0.157	0.220	−0.062
	(0.074)	(0.061)	(0.035)
Log Land Price	−0.019	−0.028	0.009
	(0.032)	(0.026)	(0.015)
Log Average Female Wage	−0.193	−0.132	−0.060
	(0.210)	(0.173)	(0.099)
Num. obs.	80	80	80

Sources: Standard errors are in parenthesis. All explanatory variables are measured in 2000 unless otherwise noted. Labor force participation rate for women aged 20-64 is from the Census. The total fertility rate is from Vital Statistics. The financial capability index is from Table for Financial Capability Indices of Prefectures constructed by Ministry of Internal Affairs and Communications. The land price is the average land price per square meter in residential areas, which is taken from Survey on Land Price of Prefectures by Ministry of Land, Infrastructure, Transport and Tourism. The mean female wage is calculated by dividing scheduled cash earnings by scheduled hours of work, which are from Basic Survey of Wage Structure 2001. For data consistency, we omit City of Yokosuka and non-major cities in the Province of Kanagawa, although they are included in the main analysis.

The second column shows how changes in child population are correlated with regional char-

[11] According to School Basic Survey 2011 conducted by the Ministry of Education, about a half of high-school graduates in Aomori, Iwate, and Akita (smaller provinces in the North East) find a job outside their home provinces, while only 11.9% of high-school graduates in Tokyo do so.

acteristics in 2000. Note that a positive coefficient implies child population increases with the variable of interest. The female labor force participation rate is negatively correlated with the growth of child population. Given the conflict between work and family life, a higher labor force participation rate may lead to a lower fertility rate. The financial capability index is positively correlated with the growth of child population. This may reflect the geographic mobility of young adults to large cities, because the financial status of major cities are generally better than that of smaller cities. We also note that a better financial status may increase the fertility rate by providing a better support for young families.

The third column shows determinants of changes in the coverage rate. The coefficient of each variable is given by subtracting the corresponding coefficient for the population factor from that for the supply factor (see Equation 13). Because the supply and population factors offset each other, the coefficients are small, although the coefficient for the female labor force participation rate is marginally significant at the 10% level. Although the supply of childcare or child population is not random, the change of the coverage rate is only weakly correlated with regional characteristics. In the following, we account for a possible policy endogeneity by including the interaction of these regional characteristics and year dummies, although doing so has little influence on our results.

B Treatment Parameters (For Online Publication)

B.1 Weights for Aggregate Treatment Parameters

We calculate treatment parameters following the method outlined by Cornelissen et al. (2015); Cornelissen, Dustmann, Raute, and Schönberg (2016). Let x_i and p_i be a vector of control variables and the propensity score for family i. The unobserved component of the MTE is denoted by $K^{'}(u_D)$. The sample mean of the propensity score is $\bar{p} = 1/N\sum_{i=1}^{N} p_i$. The ATE, TT, and TUT are given by

$$\text{ATE} \;=\; \frac{1}{N}\sum_{i=1}^{N} x_i(\beta_1 - \beta_0) + \int_0^1 K^{'}(u)du$$

$$
\begin{aligned}
\text{TT} &= \frac{1}{N}\sum_{i=1}^{N}\frac{p_i}{\bar{p}}x_i(\beta_1-\beta_0)+\int_0^1 K^{'}(u)\cdot\frac{1/N\sum_{i=1}^{N}I(p_i>u)}{\bar{p}}du \\
\text{TUT} &= \frac{1}{N}\sum_{i=1}^{N}\frac{1-p_i}{1-\bar{p}}x_i(\beta_1-\beta_0)+\int_0^1 K^{'}(u)\cdot\frac{1/N\sum_{i=1}^{N}I(p_i\leq u)}{1-\bar{p}}du.
\end{aligned}
$$

The integral can be easily calculated by discretizing the grid for u_D.

Denote the propensity score under the baseline policy by p_i and the propensity score under the alternative policy by $p_i^{'}$. The sample means of the propensity scores under these two policies are denoted by $\bar{p}$ and $\bar{p}^{'}$. The policy-relevant treatment effect (PRTE) is given by

$$
\text{PRTE} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_i^{'}-p_i}{\bar{p}^{'}-\bar{p}}x_i(\beta_1-\beta_0)+\int_0^1 K^{'}(u)\cdot\frac{1/N\sum_{i=1}^{N}I(p_i^{'}>u)-1/N\sum_{i=1}^{N}I(p_i>u)}{\bar{p}^{'}-\bar{p}}du.
$$

B.2 Estimates for Aggregate Treatment Parameters

The conventional treatment effect parameters can be calculated by aggregating MTE with proper weights. The weights for ATE is uniform, and those for TT and TUT are graphically presented in Figure 3. For TT, individuals with lower values of unobserved resistance are given more weights, while for TUT, individuals with higher values of unobserved resistance are given more weights.

Table 10 reports ATE, TT, TUT, and the difference between TT and TUT. ATEs are similar to our baseline IV estimates (see Table 7). TUTs tend to be stronger (or "better") than TT on mother's labor supply and child outcomes. For childcare expenses, TT is significantly larger than TUT, while TUT is significantly larger than TT for other expenses than childcare. No significant differences are found for other outcomes.

C Additional Tables and Figures (For Online Publication)

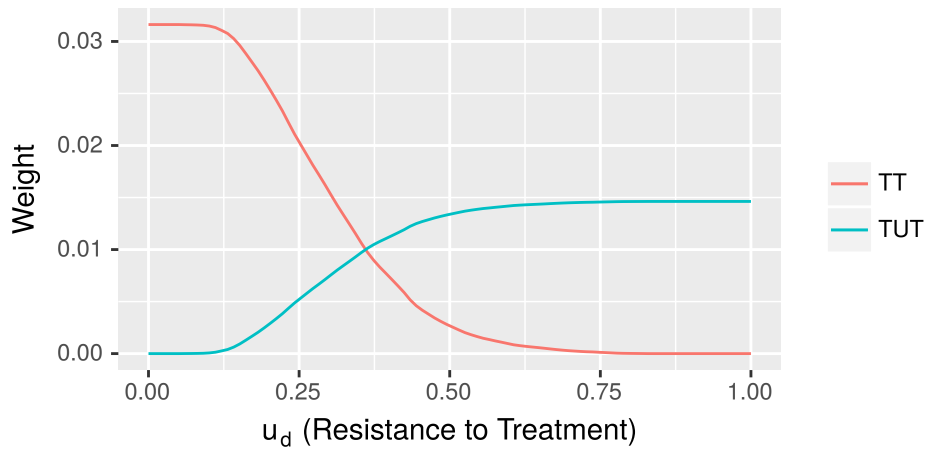

Figure 3: Weights for TT and TUT

Table 10: Aggregate Treatment Effect Parameters

	ATE	TT	TUT	TT - TUT
Market Participation	0.601	0.523	0.637	-0.114
	(0.097)	(0.116)	(0.115)	(0.133)
Hours Worked	27.949	17.544	32.766	-15.223
	(3.973)	(5.024)	(4.556)	(5.294)
Language Development	0.614	0.347	0.737	-0.390
	(0.192)	(0.300)	(0.215)	(0.327)
Aggression	-0.324	0.311	-0.615	0.926
	(0.308)	(0.424)	(0.317)	(0.388)
ADHD	-0.080	0.462	-0.327	0.789
	(0.243)	(0.364)	(0.249)	(0.343)
Parenting Quality	0.008	0.138	-0.052	0.189
	(0.226)	(0.351)	(0.256)	(0.390)
Lack of Parenting Knowledge	0.044	-0.014	0.071	-0.085
	(0.075)	(0.101)	(0.081)	(0.098)
Stress	-0.303	-0.263	-0.322	0.059
	(0.284)	(0.404)	(0.261)	(0.288)
Subjective Well-Being	0.003	0.181	-0.079	0.260
	(0.189)	(0.244)	(0.224)	(0.283)
Childcare Expenses	2.334	3.479	1.805	1.675
	(0.470)	(0.667)	(0.493)	(0.639)
Other Expenses	0.863	0.136	1.200	-1.064
	(0.519)	(0.609)	(0.610)	(0.673)

Note: Standard errors are clustered at the region level. ATE is the average treatment effect, TT is the treatment effect on the treated, and TUT is the treatment effect on the untreated.

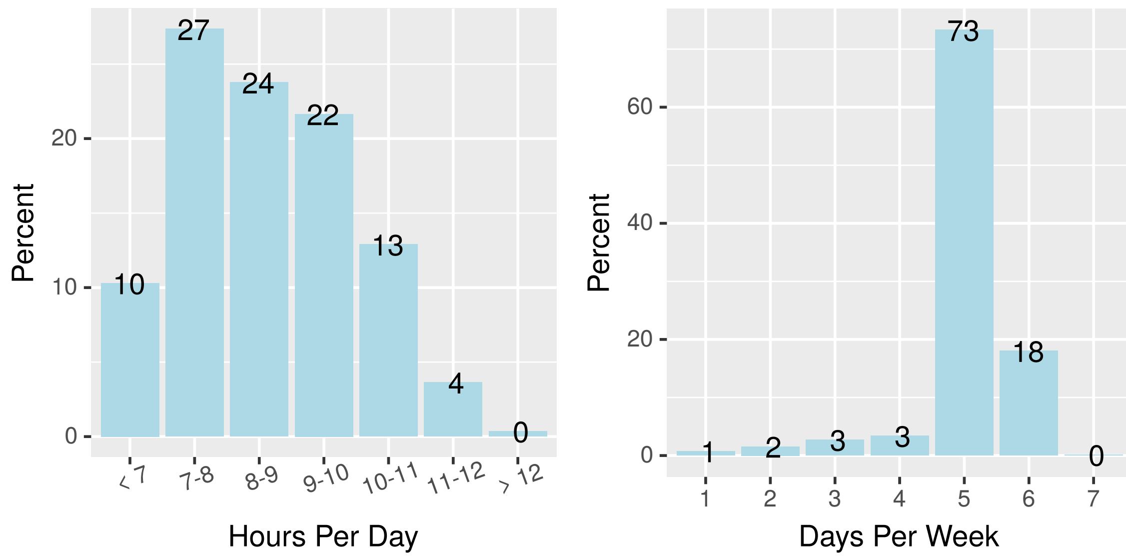

Figure 4: Distributions of Hours and Days in Childcare Center

Source: LSN21.

Note: All children are in two-parent family and at age 2½.

Table 11: Coordinates of Each Response

	Explain	Just No	Ignore	Spank	Confine
Always	0.358	−2.160	−3.422	−4.480	−7.311
Sometimes	−1.802	0.416	−0.165	−0.850	−1.592
Never	−1.828	1.489	1.271	0.471	0.438

Source: LSN21 and authors' calculation.

Note: The coordinates of each response in the multiple correspondence analysis are reported. The number in the cell indicates how each item increases/decreases the parenting quality index.

Table 12: Reduced-Form Regressions for Mother's Labor Supply and Child Outcomes

	Mother's Labor Supply		Child Outcomes		
	Work	Hours	Language	ADHD	Aggression
Model 1					
Coverage Rate	0.562	25.287	0.381	-0.025	0.490
	(0.131)	(5.894)	(0.234)	(0.414)	(0.301)
Model 2					
Propensity Score	0.612	25.873	0.559	-0.295	-0.048
	(0.109)	(4.273)	(0.170)	(0.256)	(0.198)

Note: Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity.

Table 13: Reduced-Form Regressions for Parents' Outcomes

	Parenting Quality	Insufficient Parenting Knowledge	Stress	Subjective Well-Being	Childcare Expenses	Non-Childcare Expenses
Model 1						
Coverage Rate	-0.341	-0.021	-0.166	0.095	1.736	0.120
	(0.303)	(0.082)	(0.243)	(0.242)	(0.605)	(0.542)
Model 2						
Propensity Score	0.106	0.020	-0.415	0.092	2.226	1.002
	(0.188)	(0.054)	(0.218)	(0.187)	(0.453)	(0.468)

Note: Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity.

Table 14: Treatment Effect on Mother's Labor Supply and Child Outcomes by Mother's Education

	Mother's Labor Supply		**Child Outcomes**		
	Work	Hours	Language	ADHD	Aggression
Less Than HS	0.500	23.058	0.598	-0.576	-0.430
	(0.110)	(3.773)	(0.278)	(0.363)	(0.337)
High School	0.654	27.324	0.614	-0.252	0.026
	(0.084)	(3.108)	(0.191)	(0.264)	(0.207)
2-Yr College	0.689	31.124	0.579	-0.286	-0.094
	(0.083)	(3.235)	(0.194)	(0.267)	(0.201)
4-Yr University	0.573	29.111	0.686	-0.198	0.008
	(0.086)	(3.584)	(0.216)	(0.306)	(0.232)

Note: Parameter estimates for Equation (1). Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity.

Table 15: Treatment Effect on Parents' Outcomes by Mother's Education

	Parenting Quality	Insufficient Parenting Knowledge	Stress	Subjective Well-Being	Childcare Expenses	Non-Childcare Expenses
Less Than HS	0.579	-0.050	-0.628	0.542	2.240	1.471
	(0.298)	(0.073)	(0.314)	(0.277)	(0.473)	(0.649)
High School	0.049	0.040	-0.416	0.062	2.409	1.085
	(0.214)	(0.059)	(0.245)	(0.203)	(0.378)	(0.528)
2-Yr College	-0.011	0.036	-0.289	-0.056	2.743	0.730
	(0.205)	(0.063)	(0.242)	(0.197)	(0.386)	(0.504)
4-Yr University	-0.115	0.099	-0.137	-0.007	2.980	0.812
	(0.273)	(0.074)	(0.258)	(0.235)	(0.492)	(0.597)

Note: Parameter estimates for Equation (1). Standard errors are in parenthesis and clustered at the region level. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity.

Table 16: Parameter Estimates for Selection Equation

	Estimate	Std. Error
Coverage Rate and Intercept		
Intercept	0.839	0.740
Coverage Rate	2.345	1.591
Coverage Rate Squared	0.759	1.641
Mother		
Age	-0.006	0.029
Age-Sq.	0.047	0.042
Less Than HS	-1.452	0.469
HS	-1.239	0.203
2-Yr College	-0.834	0.198
Cov. Rate × Age	-0.054	0.024
Cov. Rate × Less Than HS	4.736	2.530
Cov. Rate × HS	2.993	1.092
Cov. Rate × 2-Yr College	2.101	1.106
Cov. Rate Sq. × Less Than HS	-4.803	3.176
Cov. Rate Sq. × HS	-2.128	1.269
Cov. Rate Sq. × 2-Yr College	-1.614	1.381
Father		
Age	-0.078	0.020
Age-Sq.	0.077	0.026
Less Than HS	0.810	0.123
HS	0.375	0.061
2-Yr College	0.219	0.085
Cov. Rate × Age	0.054	0.015
Cov. Rate × Less Than HS	-0.268	0.372
Cov. Rate × HS	0.072	0.178
Cov. Rate × 2-Yr College	0.442	0.263
Child		
Born in July 2001	-0.082	0.025
Born in 2010	-0.678	1.106
Normal-Birth-Weight Boy	0.062	0.013
Low-Birth-Weight Boy	0.062	0.041
Low-Birth-Weight Girl	-0.066	0.041
Region		
Local Unemployment Rate	-0.013	0.052
Child-Teacher Ratio	-0.140	0.045
Born in 2010 × Female LFP Rate	-0.127	0.764
Born in 2010 × Fertility Rate	-0.314	0.319
Born in 2010 × Financial Status	-0.260	0.226
Born in 2010 × Log Land Price	0.127	0.089
Born in 2010 × Log Mean Female Wage	0.238	0.578

Note: Parameter estimates for the logit model for selection into childcare enrollment (see Equation 7). Standard errors are in parenthesis. The control variables include the coverage rate up to the third order polynomial, ages and education of parents, sex and birth weight of children, province-level unemployment rates, region-level child-teacher ratio, and dummies for year and region. We also include interactions of pre-reform regional characteristics in 2000 and year dummies to account for the possible policy endogeneity. The coverage rate is interacted with characteristics of parents

Table 17: Parameter Estimates for Outcome Equations

	Mother's Labor Supply		**Child Outcomes**			**Parents' Outcomes**					
	Work	Hours	Language	Aggression	ADHD	Parenting Quality	Insufficient Parenting Knowledge	Stress	Subjective Well-Being	Childcare Expenses	Non-Childcare Expenses
Less Than HS	-0.053	-2.991	-0.039	0.311	0.372	-0.470	0.035	0.149	-0.366	-0.080	-0.114
	(0.030)	(1.306)	(0.062)	(0.068)	(0.066)	(0.084)	(0.022)	(0.071)	(0.061)	(0.126)	(0.148)
HS	-0.107	-4.303	-0.049	0.169	0.154	-0.283	0.023	0.027	-0.162	-0.133	-0.246
	(0.026)	(0.952)	(0.049)	(0.080)	(0.061)	(0.068)	(0.017)	(0.048)	(0.041)	(0.112)	(0.106)
2-Yr College	-0.092	-3.478	-0.002	0.097	0.099	-0.141	0.024	-0.000	-0.006	-0.167	-0.077
	(0.020)	(0.670)	(0.035)	(0.062)	(0.057)	(0.050)	(0.013)	(0.041)	(0.039)	(0.093)	(0.097)
p (propensity score)	0.362	8.697	0.202	0.824	0.890	0.107	-0.007	-0.053	0.281	4.429	-0.492
	(0.174)	(7.068)	(0.439)	(0.597)	(0.523)	(0.530)	(0.139)	(0.506)	(0.344)	(0.963)	(0.810)
p-squared	0.121	16.250	0.416	-0.988	-0.842	-0.202	0.091	-0.063	-0.278	-1.788	1.135
	(0.143)	(5.693)	(0.347)	(0.409)	(0.363)	(0.415)	(0.104)	(0.307)	(0.302)	(0.672)	(0.717)
$p\times$ Less Than HS	-0.067	-5.021	-0.046	-0.449	-0.505	0.577	-0.134	-0.460	0.494	-0.894	0.659
	(0.088)	(3.705)	(0.165)	(0.159)	(0.197)	(0.196)	(0.052)	(0.182)	(0.169)	(0.339)	(0.427)
$p\times$ HS	0.135	2.170	0.017	-0.189	-0.093	0.136	-0.040	-0.288	0.033	-0.610	0.476
	(0.065)	(2.131)	(0.109)	(0.147)	(0.115)	(0.130)	(0.030)	(0.083)	(0.091)	(0.239)	(0.266)
$p\times$ 2-Yr College	0.178	5.898	-0.019	-0.189	-0.182	0.081	-0.049	-0.171	-0.074	-0.160	0.083
	(0.049)	(1.713)	(0.085)	(0.123)	(0.124)	(0.106)	(0.028)	(0.092)	(0.099)	(0.217)	(0.250)

Note: Parameter estimates for Equation (10). Standard errors are in parenthesis and clustered at the region level. Linear and quadratic terms of the propensity score and the interaction of the propensity score and mother's education dummies are in the outcome equation. The control variables include ages and education of parents, sex and birth weight of children, the province-level unemployment rate, the region level child-teacher ratio, and dummies for year and region. We also include interactions of the pre-reform regional characteristics in 2000 and year dummies to address the possible policy endogeneity.

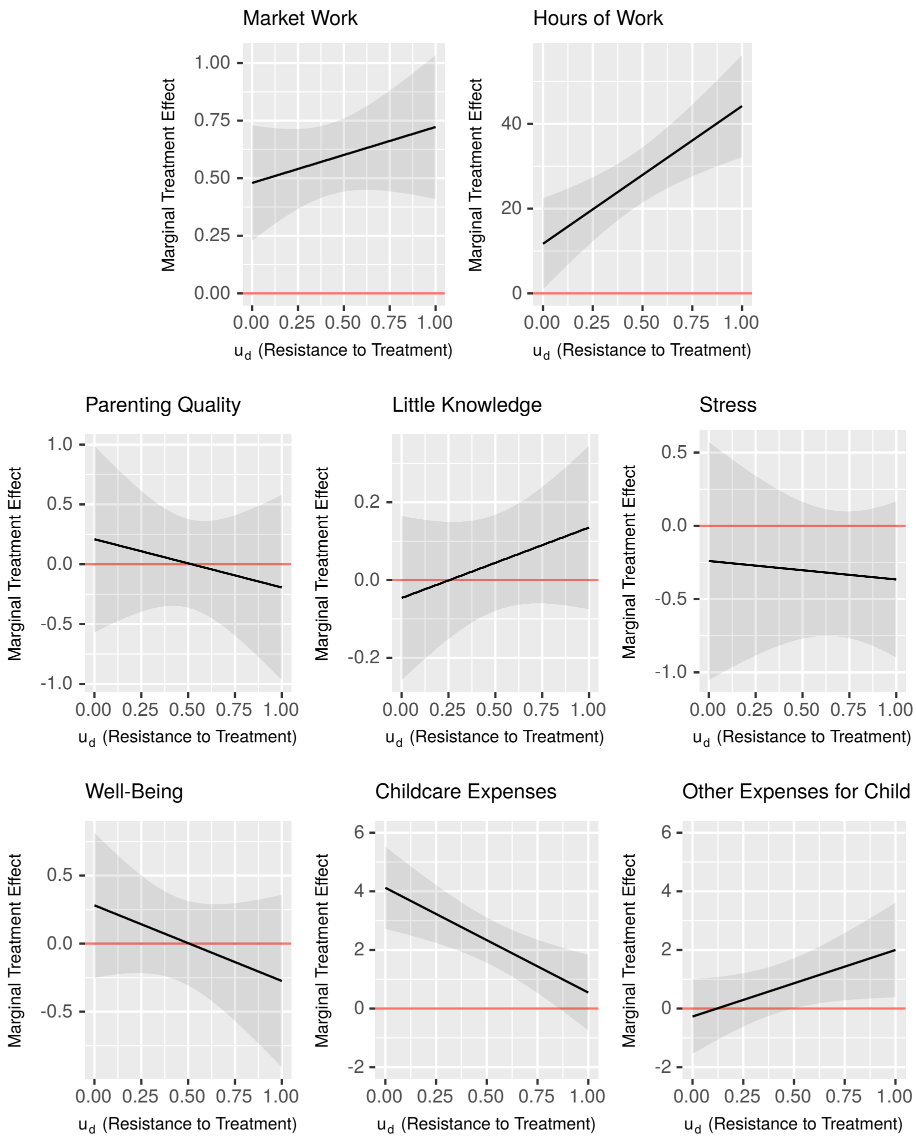

Figure 5: Marginal Treatment Effect Curves for Parents' Outcomes

Note: The marginal treatment effect is graphically presented with the 90% confidence interval. The standard errors are clustered at the region level. The MTE curve is based on the estimated outcome equation (10) and evaluated at the mean of all covariates except for the propensity score.