

Discussion Paper Series A No.656

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

Shintaro Yamaguchi

(Department of Economics, McMaster University)

Yukiko Asai

(Institute of Social Science, the University of Tokyo)

and

Ryo Kambayashi

(Institute of Economic Research, Hitotsubashi University)

March 2017

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan

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

Shintaro Yamaguchi[†] Yukiko Asai[‡] Ryo Kambayashi[§]

March 14, 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 among boys and reduces aggression and the symptoms of ADHD among the children of low-education mothers. Estimates show that the improved child behavior is strongly associated with better parenting quality and maternal wellbeing. Evidence also suggests that promoting positive parenting practices is an important element of an effective childcare program. 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 California, Berkeley, the University of Toronto, and the University of British Columbia for their helpful comments. All remaining errors are our own. Yamaguchi gratefully acknowledges financial support from the Murata Science Foundation. Asai appreciates a Grant-in-Aid for Young Scientists (B) 15K17071 and Kambayashi is supported by a Japan Society for the Promotion of Science KAKENHI Grant No. 16H02020. The data in the Longitudinal Survey of Newborns in the 21st Century were generously provided by the Japanese Ministry of Health, Labour and Welfare.

[†]Department of Economics, McMaster University, 1280 Main St. W., Hamilton, ON. Canada L8S 4M4. Email: yamtarom@mcmaster.ca.

[‡]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.

Advocates of early childhood education often cite the success of model programs, such as the Perry Preschool Program in the US, targeted at economically disadvantaged children. Evidence from these programs shows that program participation improves not only the cognitive skills but also the socioemotional skills of children. In fact, further examination reveals that the effects on cognitive skills quickly fade, but those on socioemotional skills can persist over many years. Indeed, Heckman, Pinto, and Savelyev (2013) conclude that the program leads to less crime, lower receipt of welfare, and greater voter turnout through improving socioemotional skills. Overall, their findings suggest that the preschool program has positive externalities and thereby benefits the whole of society, which helps justify the policy intervention.

While the success of model programs is appealing, it is not immediately obvious whether the results are generalizable to a large-scale universal childcare program. For instance, childcare programs may not benefit children from middle- and high-income families because they generally have better home environments and/or access to alternative high-quality childcare. In addition, maintaining the quality of childcare becomes challenging as the program scale increases.

To address this issue, an increasing number of studies estimate the effects of a universal childcare program on cognitive test scores. However, only a few studies consider the effects on socioemotional skills. Baker, Gruber, and Milligan (2008) find that a universal childcare program in Quebec actually harms children's socioemotional skills. In a subsequent study, Baker, Gruber, and Milligan (2015) show that these negative effects on socioemotional skills persist and can lead to poorer health, lower life satisfaction, and higher crime rates later in life. However, while their analysis is compelling, such negative effects are not found in other countries.

For example, Datta Gupta and Simonsen (2010) evaluate the universal childcare program in Denmark and identify no significant effects on noncognitive child outcomes. Likewise, Berlinski, Galiani, and Gertler (2009) and Felfe and Lalive (2015) estimate the effects on children's behavioral outcomes in Argentina and Germany, respectively, and find childcare enrollment improves children's behavior. Accordingly, the existing evidence concerning the effect of childcare on socioemotional skills is mixed. Furthermore, it is not entirely clear exactly how childcare enrollment affects children's socioemotional skills.

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. Specifically, we examine how childcare enrollment changes the quality of parenting, the amount of child-related expenses outside of childcare, parents' knowledge of good parenting practices, and parental wellbeing, all of which may eventually also affect the children.

Another key feature of this paper is that we estimate richer treatment effect heterogeneity than most previous studies by adopting the marginal treatment effect (MTE) framework (see Björklund and Moffitt (1987) and Heckman and Vytlacil (2005)). The MTE is a treatment effect varying by the unobserved propensity for childcare use, and this framework enables us to identify who would benefit most from childcare and how likely they are to use it. Through counterfactual simulation, we assess how children with different treatment effects are expected to be gradually induced into treatment as childcare reform progresses.

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. These treatment effects are stronger for boys, particularly low-birthweight boys. Because the language development of boys is slower than that of girls if not enrolled, childcare enrollment addresses the gap between them.

The estimates consistently indicate that childcare enrollment reduces aggression and the symptoms of attention deficit hyperactivity disorder (ADHD) among children of low-education mothers relative to those of university-educated mothers. We also find childcare enrollment tends to improve the behavior of boys and low-birthweight children, but the estimates are sometimes noisy. These children demonstrate greater aggression and ADHD symptoms 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 also examine how their home environment changes in response to childcare enrollment. The estimates show that childcare enrollment improves monetary investment in children, parenting quality, and the wellbeing of low-education mothers. Further analysis suggests that a more positive home environment for children is brought about by informing low-education mothers about better parenting practices, 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, our estimates for the MTE 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 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. Studies in Argentina (Berlinski et al. (2009)), Germany (Cornelissen, Dustmann, Raute, and Schönberg (2015) and Felfe and Lalive (2015)), Norway (Havnes and Mogstad (2011, 2015) and Drange and Havnes (2015)), (Felfe, Nollenberger, and Rodríguez-Planas (2015)), and the US (Cascio and Schanzenbach (2013)) all point to gains in test scores from participation in early childhood education, and some identify stronger effects for disadvantaged children. 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 et al. (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 most benefits children from disadvantaged families, but 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 et al. (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 impacts upon 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. 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.

Finally, most extant studies account for heterogeneity in treatment effects only using the observed characteristics. Cornelissen et al. (2015) and Felfe and Lalive (2015) are exceptions and estimate the MTE that varies by the unobserved characteristics of German children. 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. While our results echo these findings, 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 located¹. 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 other-

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

wise noted².

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 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. Figure 4 in Appendix C provides detailed statistics.

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.

²The 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.

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.

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 dif-

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

⁴For 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.

ferred considerably across regions. We exploit this variation in the speed of childcare expansion between regions to identify the causal effects. We extensively discuss our identification strategy in Section 5.3.

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 Canada⁵.

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 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 7 in Appendix A for details.

⁵In 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.

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 treatment status at other ages, the treatment and control groups exhibit very different childcare enrollment patterns over time. At 6 months of age, few children are enrolled in childcare regardless of treatment status 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 enrollment is serially correlated and that the enrollment patterns over time are very different for the treated and the untreated. The difference in enrollment patterns between the treated and untreated would be smaller if treatment was defined by enrollment at ages other than 2½ years.

4.2.2 Child Outcomes

We construct measures according to child language development, aggression, and ADHD symptoms. These measures are constructed from a set of questions that can be answered with a simple 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 9 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 subjective wellbeing is measured by the question “What makes you happy when you raise your child?” The following nine items are listed and the respondents select all that apply: “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”.

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 1 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 0.372, 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

Table 1: 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
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
Lack of Parenting Knowledge	67728	0.089	0.284	0.084	0.091	0.002
Subjective Well-Being	67873	0.000	1.000	0.027	-0.012	0.000

Source: LSN21

Note: Children are in two-parent family. Child outcomes, parenting quality, and parent's 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 3.5 year old. Other variables are evaluated when children are 2.5 year old. Expenses in the survey month are measured in 10,000 JPY (\approx 100 USD).

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).

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, better subjective wellbeing, and more parenting knowledge, 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 (see Kline and Walters (2016)). Table 2 provides the distribution of childcare modes by the mothers' labor market status.

Table 2: 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 2.5 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.

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.

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. We discuss the validity of the instruments in the final subsection.

5.1 Instrumental Variable Regression

Our basic specification is based on IV regression. Define D as an indicator variable for childcare enrollment 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 characteristics. The estimation equation is given by

$$Y = \beta_{k=1} + \beta_{k=2}X_{k=2} + \cdots + \beta_{k=K}X_{k=K} + \tau_{k=1}D + \tau_{k=2}D(X_{k=2} - \bar{X}_{k=2}) + \cdots + \tau_{k=K}D(X_{k=K} - \bar{X}_{k=K}) + \varepsilon. \quad (1)$$

where ε is an error term that may be correlated with treatment status D , k indexes the elements in the vectors, and \bar{X}_k is the mean of X_k . The coefficient for childcare enrollment $\tau_{k=1}$ is the average treatment effect (ATE), because we demean the covariates X . Demeaning does not affect the interpretation of the other coefficients $\tau_{k \neq 1}$.

Childcare enrollment is determined by the following selection equation:

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

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. The validity of the instrument is extensively discussed in Section 5.3. 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 , $P(X, Z)$, X , and $P(X, Z) \cdot (X - \bar{X})$ instead of Z 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. A potential outcome Y_j for enrollment status j is given by

$$Y_j = X\beta_j + U_j \quad (3)$$

$$E(U_j|X) = 0. \quad (4)$$

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

$$D = 1\{X\gamma + \delta Z > V\} \quad (5)$$

$$= 1\{F_V(X\gamma + \delta Z) > F_V(V)\} \quad (6)$$

$$= 1\{P(X\gamma + \delta Z) > U_D\}, \quad (7)$$

where F_V is the cumulative distribution function for V , $P(\cdot)$ is the propensity score, and U_D is the quantile of V . We assume that (U_j, U_D) is independent of Z given X . We refer to U_D as the 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

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

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,

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

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

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

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,

$$MTE(X = x, U_D = p) = \frac{\partial E(Y | X = x, P(X, Z) = p)}{\partial p} \quad (11)$$

$$= x(\beta_1 - \beta_0) + \frac{\partial K(p)}{\partial p}. \quad (12)$$

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. But what are they? Given the work requirement and the rationing rule that favors fulltime workers, the mother's labor market attachment and 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 .

5.3 Threats to Identification

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. Note that this identification strategy is more robust than the approach that employs spatial variation, which is typically adopted in the MTE literature on returns to schooling.

Because we control for time and region fixed effects, the identifying variation in our approach is the regional variation in the growth of the childcare coverage rate. As discussed in Section 3.2, the financial condition of the local government and other regional factors are likely to affect the growth of the coverage rate. For example, the Cabinet Office (2010) argues that three factors can slow rollout of the childcare reform. First, the bureaucratic 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 anal-

ysis 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.

There could be some concern that 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 reform⁶. Indeed, the regulation for quality control is partly responsible for the slow progress of the childcare reform relative to other countries.

Another threat to the identification is 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 2012⁷, 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 either a major city or the rest of a province excluding its major cities, at most 4.6 percent of the sample moved between regions

⁶Local 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.

⁷Conducted by the Statistics Bureau every 5 years and covering about 1 percent of the population.

for childcare purposes. As we show in Section 6.3, selective migration appears to have little effect on the estimates.

Other issues that could affect our estimates 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.

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, and dummies for year and region. The interactions between the coverage rate and parent and child 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 Table 10 in Appendix C.

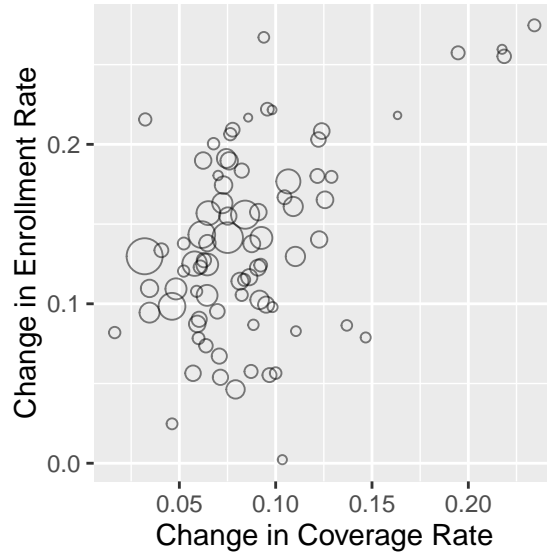


Figure 1: Growth of Coverage Rate and Enrollment Rate by Region

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.

We focus on the selected variables and report their average marginal effects in Table 3. The average marginal effect of the coverage rate is significantly positive at 0.795. The high statistical significance gives us confidence in the validity of our IV regression, but a formal test for weak IV is in the following subsection.

The mother's age has a significant 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.

Child characteristics do not have large effects on childcare enrollment. The reference group is normal-birthweight girls. The probability of childcare enrollment is only about one percentage point higher for normal-birthweight boys. The childcare enrollment rates of low-birthweight boys and girls are not significantly different from those of normal-birthweight girls.

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-statistics for testing weak instruments are well above 10 in all models, which implies that we can reject the null hypothesis that the instruments are weak. All the reported standard errors are clustered at the region level. For comparison purposes, we also report ordinary least squares (OLS) estimates in Tables 11 and 12 in Appendix C.

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. As shown in Table 4, the estimated ATEs on the mother's labor market participation and weekly hours of work are positive and significant at 0.690 and 32.128 hours, respectively.

The treatment effects vary by the mother's education. In the following, we refer to mothers with less than a high school education as low-education mothers and those with a 4-year university education or higher as high-education mothers. The reference group is high-education mothers. The estimates show that the treatment effects are smaller for low-education mothers, and the difference for hours of work is statistically significant. We do not find statistically significant differences

Table 3: First-Stage Regression: Average Marginal Effects on Childcare Enrollment at Age 2.5

	Ave. Marginal Effect
Region	
Coverage Rate	0.795 (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)
Child	
Normal-Birth-Weight Boy	0.012 (0.004)
Low-Birth-Weight Boy	0.013 (0.009)
Normal-Birth-Weight Girl	-0.013 (0.008)

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 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 and children. Parameter estimates for the logit model are reported in Table 10 in Appendix C.

in treatment effects by the child's sex or birthweight.

6.2.2 Child Development and Behavior

We next provide estimates for child development and behavior in Table 4. To facilitate interpretation, all of these outcome variables are normalized to have a mean of zero and a standard deviation of one.

The ATE on language development is positive and significant. Language development is significantly slower for boys, particularly those with a low birthweight, when not enrolled in childcare. However, childcare enrollment narrows the lag behind normal-birthweight girls. The estimated treatment effects are 0.217 and 0.774 standard deviations larger than those for normal-birthweight girls. We find no significant difference in treatment effects by the mother's education.

The ATEs on children's aggression and ADHD symptoms are insignificant, but treatment seems effective for some groups. Without treatment, children of low-education mothers show more aggressive behavior and ADHD symptoms than the children of high-education mothers. Boys and low-birthweight girls also behave worse than normal-birthweight girls without treatment. Childcare enrollment tends to reduce aggression and ADHD symptoms among these children, although some estimates are noisy. The results for the children of low-education mothers are most consistent: the differences from high-education mothers in the treatment effects are statistically significant at -0.463 and -0.372 standard deviations for aggression and ADHD symptoms, respectively. These estimates suggest that the children of low-education mothers behave as well as the children of high-education mothers when enrolled in childcare.

6.2.3 Parents' Outcomes

To better understand the mechanism behind how childcare enrollment affects child outcomes, we estimate the effects on the children's home environment. Table 4 presents the IV estimates for the parenting quality index, which is normalized to have a mean of zero and a standard deviation of one.

The ATE of childcare enrollment on parenting quality is nearly zero and in-

Table 4: Effects of Childcare Enrollment on Mother's Work and Child Outcomes

	Mother's Labor Supply			Child Outcomes				Parents' Outcomes			
	Work	Hours	Language	Aggression	ADHD	Parenting Quality	Childcare Expenses	Non-Childcare Expenses	Insufficient Parenting Knowledge	Subjective Well-Being	
Coefficients											
Low-Education Mother	-0.031 (0.021)	-1.390 (0.922)	-0.009 (0.055)	0.329 (0.069)	0.234 (0.065)	-0.505 (0.091)	-0.137 (0.115)	-0.006 (0.146)	0.043 (0.021)	-0.400 (0.059)	
Normal Birth Weight Boy	-0.015 (0.008)	-0.346 (0.344)	-0.387 (0.020)	0.341 (0.017)	0.238 (0.020)	-0.178 (0.019)	0.046 (0.020)	-0.061 (0.063)	0.000 (0.005)	-0.040 (0.021)	
Low Birth Weight Boy	-0.021 (0.019)	-0.057 (0.741)	-0.922 (0.089)	0.392 (0.054)	0.441 (0.079)	-0.245 (0.077)	0.033 (0.054)	-0.303 (0.170)	0.013 (0.021)	-0.118 (0.060)	
Low Birth Weight Girl	-0.010 (0.019)	0.069 (0.898)	-0.083 (0.055)	0.166 (0.056)	0.117 (0.064)	-0.105 (0.073)	0.099 (0.068)	-0.063 (0.109)	0.017 (0.021)	-0.117 (0.067)	
Childcare	0.690 (0.087)	32.128 (3.173)	0.460 (0.212)	0.095 (0.215)	-0.356 (0.295)	0.010 (0.235)	2.865 (0.423)	1.320 (0.559)	0.036 (0.064)	-0.029 (0.215)	
Childcare × Low-Edu.	-0.082 (0.065)	-6.573 (2.770)	-0.096 (0.159)	-0.463 (0.223)	-0.372 (0.180)	0.692 (0.244)	-0.780 (0.291)	0.614 (0.387)	-0.149 (0.054)	0.577 (0.170)	
Childcare × NBW Boy	0.014 (0.021)	-0.633 (1.017)	0.217 (0.059)	-0.161 (0.051)	-0.045 (0.062)	0.049 (0.067)	-0.170 (0.069)	0.013 (0.174)	0.017 (0.016)	0.086 (0.059)	
Childcare × LBW Boy	-0.006 (0.053)	-2.530 (2.394)	0.774 (0.265)	0.063 (0.163)	-0.307 (0.208)	0.119 (0.202)	-0.086 (0.145)	1.016 (0.579)	0.026 (0.058)	0.111 (0.156)	
Childcare × LBW Girl	0.004 (0.062)	-0.398 (2.845)	-0.192 (0.175)	-0.354 (0.201)	-0.160 (0.217)	0.111 (0.225)	-0.287 (0.235)	0.387 (0.386)	0.009 (0.066)	0.243 (0.228)	
Test Statistic											
Weak Instrument	15.390	15.547	15.522	13.900	14.868	14.799	15.246	15.057	15.407	15.146	

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, 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. Low-education mothers are those with less than high school education. The reference group for mother's education is mothers graduated from 4-year university or higher. NBW and LBW denote normal birth weight and low birth weight, respectively. The reference group for child's sex and birth weight is normal-birth-weight girls.

significant. However, the treatment effect for low-education mothers is 0.692 standard deviations greater than that for high-education mothers, and is statistically significant. The parenting quality of low-education mothers is 0.505 standard deviations lower than that of high-education mothers without treatment, but childcare enrollment improves it sufficiently that the gap disappears. Childcare enrollment does not seem to significantly improve the parenting quality of mothers of boys and low-birthweight children, even though their parenting quality is lower without treatment.

Childcare enrollment also affects monetary expenses for children. Note that these monetary expenses are specifically for the surveyed child, and any expenses for the child's siblings are not included. The ATE on childcare expenses is positive and significant at 28,650 JPY. The treatment effect for low-education mothers is significantly lower, by 7,800 JPY, than that for high-education mothers, which is consistent with the fact that a lower fee is charged for low-income families. The magnitudes of the differences in treatment effects are not large for the other subgroups.

The ATE on nonchildcare expenses is positive and significant at 13,200 JPY, which implies that childcare and other child-related services and goods are gross complements, rather than substitutes. The treatment effects for low-education mothers, boys, and low-birthweight children are greater than for their reference groups, although the estimates are imprecise.

To see why the parenting quality of low-education mothers improves through childcare use, we estimate the treatment effects on those mothers who self-report that they do not have good parenting knowledge. The ATE is nearly zero and insignificant, but the treatment effects for low-education mothers are significantly lower, by 14.9 percentage points, than for high-education mothers, which implies that fewer low-education mothers report a lack of parenting knowledge once treated. Childcare enrollment therefore seems to inform low-education mothers about better parenting practices.

Finally, we estimate the effects on parents' subjective wellbeing. The index is normalized so that the mean is zero and the standard deviation is one. The ATE is nearly zero and insignificant, but treatment benefits parents in some subgroups.

If their children are not enrolled in childcare, low-education mothers and parents of boys and low-birthweight children report significantly lower wellbeing than the corresponding reference groups. The treatment effects for low-education mothers are 0.577 standard deviations higher than those for high-education mothers, and are statistically significant. The treatment effects for boys and low-birthweight children are also greater than those for normal-birthweight girls, but not statistically significant.

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 ATE 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 $2\frac{1}{2}$ 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 $\frac{1}{2}$ and $1\frac{1}{2}$ 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 5. 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 ATE on nonchildcare expenses is small and insignificant. When the number of siblings is included (Model 4), the ATEs 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.

6.4 Interpretation and Discussion

6.4.1 Child Outcomes

The estimates indicate that childcare enrollment on average improves the language development of children. In addition, the treatment effects are stronger for boys, particularly low-birthweight boys, than for normal-birthweight girls, such that they perform almost equally well when enrolled in childcare.

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)).

Table 5: Robustness Checks

	(1) Benchmark	(2) Endogenous Fertility	(3) Selective Migration	(4) Siblings	(5) Lagged Cov. Rate	(6) Linear in 1st Stage
Ave. Treatment Effect						
Participation	0.690 (0.087)	0.701 (0.097)	0.773 (0.062)	1.114 (0.098)	0.676 (0.083)	0.660 (0.108)
Hours	32.128 (3.173)	30.029 (3.391)	28.095 (2.347)	48.286 (3.892)	29.881 (3.509)	25.966 (3.783)
Mental Development	0.460 (0.212)	0.405 (0.227)	0.229 (0.111)	0.424 (0.180)	0.293 (0.199)	0.331 (0.232)
Aggression	0.095 (0.215)	0.012 (0.209)	0.195 (0.130)	0.041 (0.222)	0.015 (0.211)	0.053 (0.240)
ADHD Symptoms	-0.356 (0.295)	-0.480 (0.291)	0.076 (0.123)	-0.245 (0.267)	-0.349 (0.292)	-0.250 (0.301)
Parenting Quality	0.010 (0.235)	0.153 (0.223)	-0.157 (0.167)	-0.209 (0.179)	-0.046 (0.222)	0.280 (0.294)
Childcare Expense	2.865 (0.423)	3.496 (0.466)	2.233 (0.274)	2.034 (0.322)	2.902 (0.407)	3.840 (0.535)
Other Expense	1.320 (0.559)	1.692 (0.629)	0.206 (0.330)	0.848 (0.431)	1.213 (0.544)	1.265 (0.654)
Lacking Parenting Knowledge	0.036 (0.064)	0.042 (0.066)	0.007 (0.036)	0.008 (0.068)	0.027 (0.063)	0.004 (0.063)
Well-Being	-0.029 (0.215)	-0.202 (0.219)	0.122 (0.129)	-0.092 (0.214)	-0.058 (0.213)	-0.115 (0.214)
Diff. in ATE for Low-Educ. Mothers						
Participation	-0.082 (0.065)	-0.088 (0.064)	-0.083 (0.066)	-0.137 (0.067)	-0.084 (0.064)	-0.087 (0.065)
Hours	-6.573 (2.770)	-6.502 (2.721)	-6.690 (2.906)	-9.718 (2.878)	-6.637 (2.706)	-6.414 (2.717)
Mental Development	-0.096 (0.159)	-0.100 (0.157)	-0.084 (0.162)	0.003 (0.137)	-0.074 (0.153)	-0.097 (0.155)
Aggression	-0.463 (0.223)	-0.462 (0.224)	-0.484 (0.210)	-0.403 (0.205)	-0.459 (0.218)	-0.420 (0.227)
ADHD Symptoms	-0.372 (0.180)	-0.388 (0.185)	-0.432 (0.179)	-0.313 (0.166)	-0.367 (0.179)	-0.370 (0.179)
Parenting Quality	0.692 (0.244)	0.707 (0.247)	0.659 (0.251)	0.636 (0.209)	0.692 (0.242)	0.725 (0.252)
Childcare Expense	-0.780 (0.291)	-0.859 (0.332)	-0.748 (0.267)	-0.650 (0.241)	-0.772 (0.293)	-0.853 (0.364)
Other Expense	0.614 (0.387)	0.591 (0.400)	0.601 (0.407)	0.724 (0.303)	0.614 (0.382)	0.594 (0.372)
Lacking Parenting Knowledge	-0.149 (0.054)	-0.151 (0.053)	-0.146 (0.050)	-0.102 (0.051)	-0.147 (0.054)	-0.153 (0.055)
Well-Being	0.577 (0.170)	0.607 (0.163)	0.560 (0.161)	0.544 (0.158)	0.571 (0.166)	0.616 (0.169)

Note: See note for Table 4 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 woman 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 0.5 and 1.5 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.

Our estimates indicate that the ATEs on child aggression and ADHD symptoms are almost zero. However, childcare enrollment significantly improves the behavior of children in some subgroups. Children of low-education mothers 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. Some of the estimated effects on the behavior of boys and low-birthweight girls are positive, but they are not consistently significant.

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 Why Do Child Outcomes Improve?

Why does childcare enrollment improve language development and the behavior of some children? There are direct and indirect paths. One is 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 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 labor market attachment and skills. 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 changes with unobserved resistance to treatment u_D . The three panels at the top present the MTE curves for child development and behavioral outcomes. 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 two panels in the middle plot the MTE curves for expenses for the surveyed child. The MTE on childcare expenses is positive and significant and decreases with unobserved resistance. This is reasonable given that the childcare fees increase with household income and that mothers with a high resistance are likely to be unskilled. The MTE on expenses other than childcare is upward sloping, and is positive and significant for higher values of unobserved resistance. These results again suggest that childcare and other goods and services for the child are gross complements.

The bottom three panels illustrate the MTE curves for parenting quality, parenting knowledge, and subjective wellbeing. The estimates are 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 Table 8 and Figure 3 in Appendix B, respectively. 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 6 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 in formal childcare in Denmark, which has the highest childcare enrollment rate in the OECD.

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 treat-

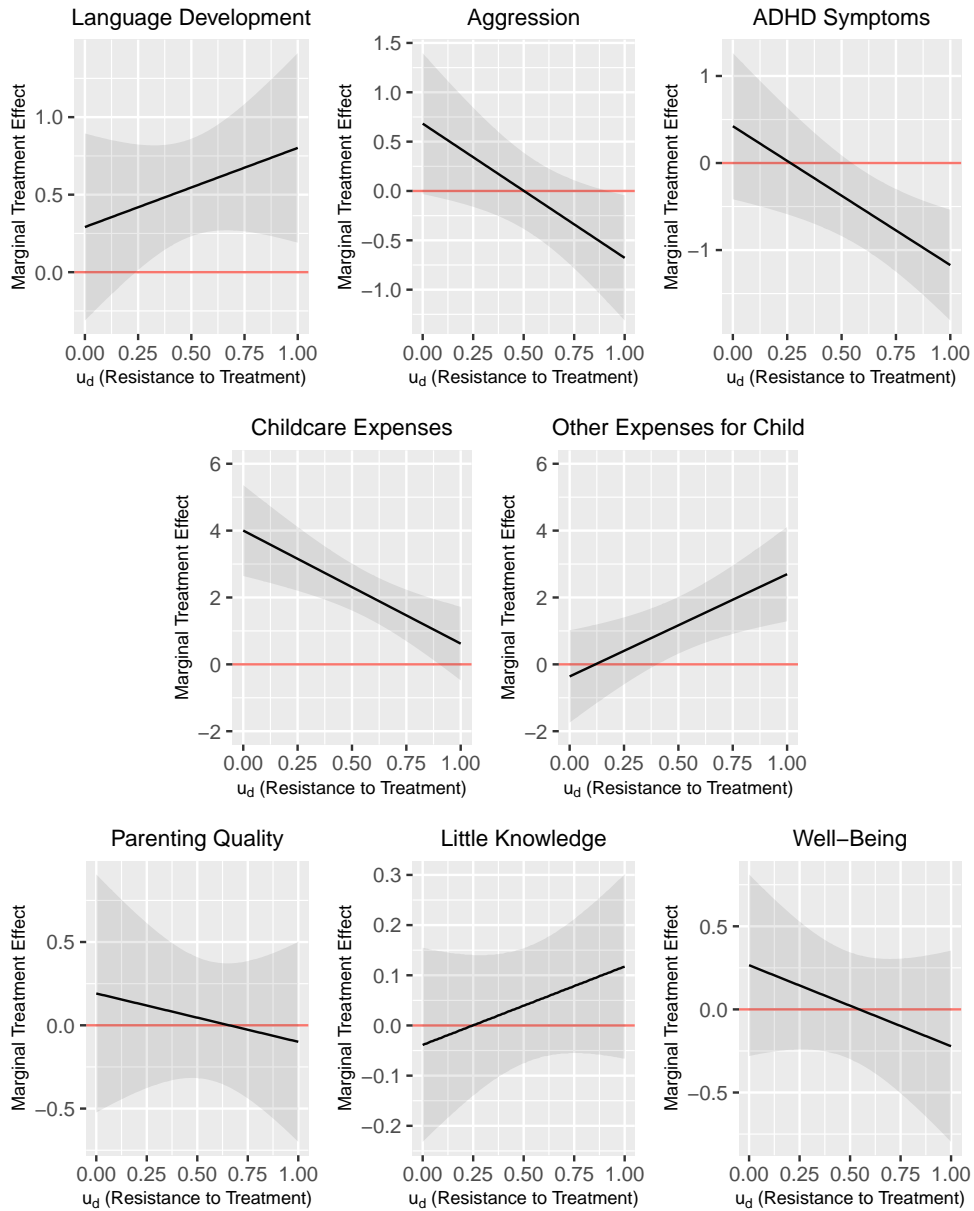


Figure 2: Marginal Treatment Effect Curves

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.

Table 6: Counterfactual Policy Simulations

	Policy-Relevant Treatment Effect			Propensity Score	
	Mental Development	Aggression	ADHD	Baseline	New Policy
Raise Coverage Rate by 0.1	0.514 (0.191)	0.099 (0.209)	-0.258 (0.301)	0.392 (0.012)	0.480 (0.018)
Raise Coverage Rate by 0.2	0.535 (0.184)	0.043 (0.201)	-0.324 (0.292)	0.392 (0.012)	0.568 (0.029)
Raise Coverage Rate by 0.3	0.554 (0.181)	-0.009 (0.197)	-0.385 (0.286)	0.392 (0.012)	0.649 (0.039)

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 Appendix B.

ment 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 some children. These children show slower language development and more aggression and ADHD symptoms than other children when they are not enrolled in childcare. However, childcare enrollment helps them catch up and provides a level playing field.

The estimates indicate that childcare enrollment is particularly effective for improving the behavior of children of low-education mothers. Part of the strong positive effect may be indirectly 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. 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 interpre-

tation 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.
- 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 childcare 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 Child-care (For Online Publication)

Cabinet Office (2010) argues that three factors slow down the rollout of the child-care 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

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

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.⁸ 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 chil-

⁸See Section 4 for our definition of region and selection criteria.

dren, 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,⁹ which may eventually affect the size of child population.

Table 7 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.

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

⁹According 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.

Table 7: 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 (0.190)	-0.743 (0.156)	0.186 (0.089)
Total Fertility Rate	0.154 (0.106)	0.156 (0.087)	-0.001 (0.050)
Financial Capability Index	0.157 (0.074)	0.220 (0.061)	-0.062 (0.035)
Log Land Price	-0.019 (0.032)	-0.028 (0.026)	0.009 (0.015)
Log Average Female Wage	-0.193 (0.210)	-0.132 (0.173)	-0.060 (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.

istics. 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

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

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 8 reports ATE, TT, TUT, and the difference between TT and TUT. ATEs are similar to our baseline IV estimates (see Table 5). 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.

Table 8: Aggregate Treatment Effect Parameters

	ATE	TT	TUT	TT - TUT
Market Participation	0.591 (0.092)	0.546 (0.118)	0.611 (0.104)	-0.065 (0.125)
Hours Worked	27.398 (3.814)	17.728 (5.169)	31.876 (4.303)	-14.148 (5.398)
Language Development	0.546 (0.191)	0.383 (0.277)	0.622 (0.214)	-0.239 (0.297)
Aggression	0.002 (0.235)	0.440 (0.339)	-0.198 (0.243)	0.638 (0.317)
ADHD	-0.375 (0.282)	0.139 (0.410)	-0.609 (0.270)	0.748 (0.336)
Parenting Quality	0.046 (0.221)	0.139 (0.337)	0.003 (0.224)	0.136 (0.315)
Childcare Expenses	2.311 (0.427)	3.395 (0.646)	1.811 (0.424)	1.584 (0.588)
Other Expenses	1.168 (0.516)	0.188 (0.675)	1.621 (0.562)	-1.433 (0.635)
Lack of Parenting Knowledge	0.039 (0.070)	-0.011 (0.094)	0.062 (0.074)	-0.073 (0.086)
Subjective Well-Being	0.022 (0.195)	0.178 (0.258)	-0.050 (0.216)	0.228 (0.262)

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.

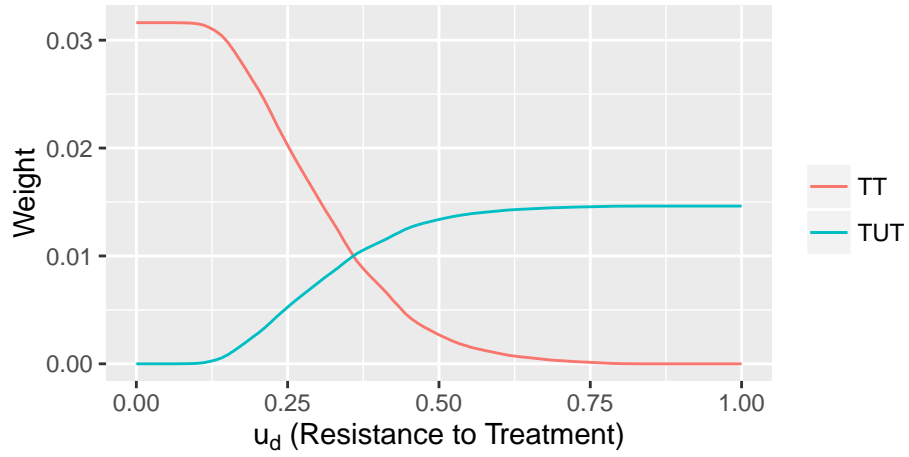


Figure 3: Weights for TT and TUT

C Additional Tables and Figures (For Online Publication)

Table 9: 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.

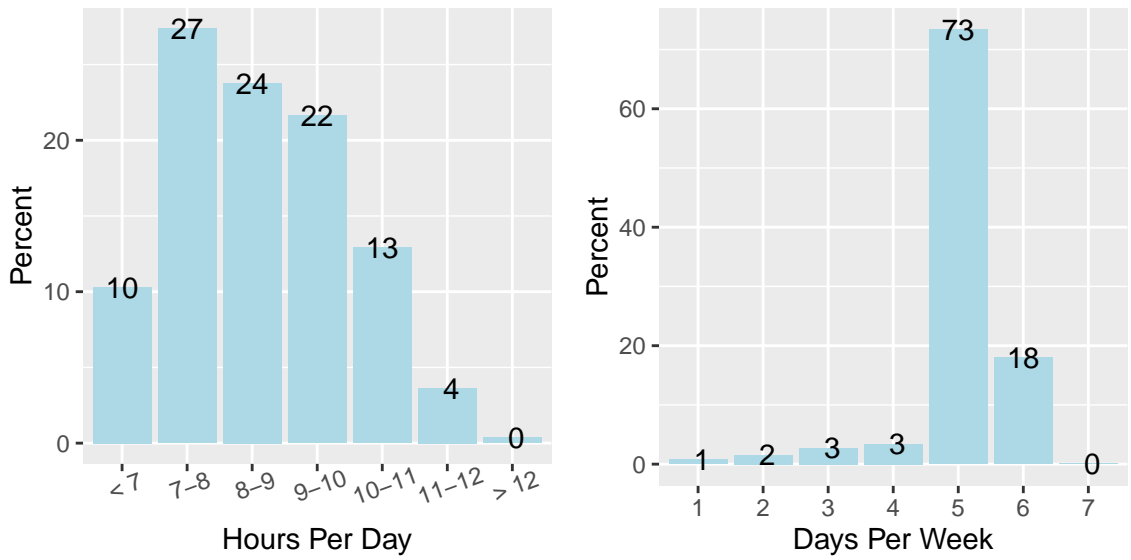


Figure 4: Distributions of Hours and Days in Childcare Center

Source: LSN21.

Note: All children are in two-parent family and at age 2.5.

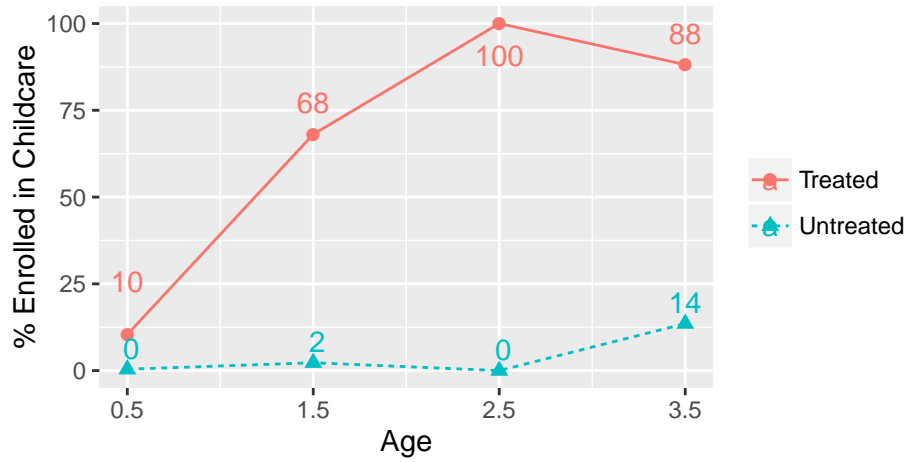


Figure 5: Childcare Enrollment Over Time

Source: LSN21

Note: Treated children are those enrolled in childcare at age 2.5. Untreated children are those not enrolled.

Table 10: Parameter Estimates for Selection Equation

	Estimate	Std. Error
Coverage Rate and Intercept		
Intercept	-0.328	0.530
Coverage Rate	2.179	1.786
Coverage Rate Squared	0.954	1.710
Mother		
Age	-0.006	0.029
Age-Sq.	0.046	0.042
Less Than HS	-1.466	0.472
HS	-1.240	0.201
2-Yr College	-0.835	0.198
Cov. Rate \times Age	-0.052	0.024
Cov. Rate \times Less Than HS	4.821	2.546
Cov. Rate \times HS	3.008	1.079
Cov. Rate \times 2-Yr College	2.114	1.105
Cov. Rate Sq. \times Less Than HS	-4.924	3.205
Cov. Rate Sq. \times HS	-2.163	1.249
Cov. Rate Sq. \times 2-Yr College	-1.636	1.383
Father		
Age	-0.078	0.020
Age-Sq.	0.077	0.026
Less Than HS	0.803	0.123
HS	0.372	0.061
2-Yr College	0.216	0.086
Cov. Rate \times Age	0.054	0.015
Cov. Rate \times Less Than HS	-0.251	0.372
Cov. Rate \times HS	0.081	0.177
Cov. Rate \times 2-Yr College	0.450	0.264
Child		
Born in July 2001	-0.082	0.025
Born in 2010	-0.338	1.082
Normal-Birth-Weight Boy	0.081	0.041
Low-Birth-Weight Boy	0.248	0.111
Low-Birth-Weight Girl	0.122	0.118
Cov. Rate Sq. \times Normal-Birth-Weight Boy	-0.057	0.122
Cov. Rate Sq. \times Low-Birth-Weight Boy	-0.559	0.325
Cov. Rate Sq. \times Low-Birth-Weight Girl	-0.565	0.349
Region		
Born in 2010 \times Female LFP Rate	0.102	0.767
Born in 2010 \times Fertility Rate	-0.219	0.380
Born in 2010 \times Financial Status	-0.149	0.294
Born in 2010 \times Log Land Price	0.076	0.095
Born in 2010 \times Log Mean Female Wage	0.293	0.619

Table 11: Effects of Childcare Enrollment on Mother's Work and Child Outcomes

Coefficients	Participation		Hours of Work		Language Development		Aggression		ADHD	
	IV		IV		IV		IV		IV	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Low-Education Mother	-0.018 (0.008)	-0.031 (0.021)	-1.202 (0.352)	-1.390 (0.922)	-0.081 (0.029)	-0.009 (0.055)	0.199 (0.039)	0.329 (0.069)	0.170 (0.031)	0.234 (0.065)
Normal Birth Weight Boy	-0.010 (0.004)	-0.015 (0.008)	-0.279 (0.118)	-0.346 (0.344)	-0.375 (0.010)	-0.387 (0.020)	0.296 (0.010)	0.341 (0.017)	0.218 (0.009)	0.238 (0.020)
Low Birth Weight Boy	-0.022 (0.008)	-0.021 (0.019)	-0.552 (0.270)	-0.057 (0.741)	-0.781 (0.034)	-0.922 (0.089)	0.410 (0.024)	0.392 (0.054)	0.350 (0.030)	0.441 (0.079)
Low Birth Weight Girl	-0.008 (0.008)	-0.010 (0.019)	-0.209 (0.285)	0.069 (0.898)	-0.172 (0.024)	-0.083 (0.055)	0.072 (0.017)	0.166 (0.056)	0.080 (0.020)	0.117 (0.064)
Childcare	0.724 (0.012)	0.690 (0.087)	27.428 (0.385)	32.128 (3.173)	0.115 (0.007)	0.460 (0.212)	0.016 (0.015)	0.095 (0.215)	-0.037 (0.016)	-0.356 (0.295)
Childcare × Low-Edu.	-0.105 (0.016)	-0.082 (0.065)	-8.437 (0.625)	-6.573 (2.770)	0.008 (0.036)	-0.096 (0.159)	-0.043 (0.066)	-0.463 (0.223)	-0.044 (0.056)	-0.372 (0.180)
Childcare × NBW Boy	-0.002 (0.005)	0.014 (0.021)	-0.654 (0.212)	-0.633 (1.017)	0.192 (0.016)	0.217 (0.059)	-0.014 (0.018)	-0.161 (0.051)	0.002 (0.016)	-0.045 (0.062)
Childcare × LBW Boy	-0.005 (0.016)	-0.006 (0.053)	-0.871 (0.716)	-2.530 (2.394)	0.357 (0.050)	0.774 (0.265)	0.009 (0.043)	0.063 (0.163)	-0.026 (0.052)	-0.307 (0.208)
Childcare × LBW Girl	0.001 (0.013)	0.004 (0.062)	0.348 (0.598)	-0.398 (2.845)	0.091 (0.034)	-0.192 (0.175)	-0.035 (0.036)	-0.354 (0.201)	-0.018 (0.041)	-0.160 (0.217)
Test Statistics										
Weak Instrument	15.390		15.547		15.522		13.900		14.868	
Wu-Hausman	1.783		1.286		2.278		2.374		1.158	

Note: Parameter estimates for Equation (1). See note for Table 4.

Table 12: Effects of Childcare Enrollment on Parent's Outcomes

	Parenting Quality		Childcare Expenses for the Child		Other Expenses for the Child		Lack of Parenting Knowledge		Subjective Well-Being	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Coefficients										
Low-Education Mother	-0.335 (0.031)	-0.505 (0.091)	0.111 (0.022)	-0.137 (0.115)	0.029 (0.057)	-0.006 (0.146)	-0.010 (0.009)	0.043 (0.021)	-0.212 (0.031)	-0.400 (0.059)
Normal Birth Weight Boy	-0.169 (0.010)	-0.178 (0.019)	0.001 (0.006)	0.046 (0.020)	-0.061 (0.023)	-0.061 (0.063)	0.007 (0.003)	0.000 (0.005)	-0.013 (0.009)	-0.040 (0.021)
Low Birth Weight Boy	-0.221 (0.032)	-0.245 (0.077)	0.032 (0.019)	0.033 (0.054)	0.001 (0.053)	-0.303 (0.170)	0.027 (0.007)	0.013 (0.021)	-0.098 (0.029)	-0.118 (0.060)
Low Birth Weight Girl	-0.085 (0.028)	-0.105 (0.073)	0.015 (0.018)	0.099 (0.068)	-0.007 (0.043)	-0.063 (0.109)	0.025 (0.008)	0.017 (0.021)	-0.059 (0.019)	-0.117 (0.067)
Childcare	-0.018 (0.014)	0.010 (0.235)	2.646 (0.052)	2.865 (0.423)	-0.016 (0.033)	1.320 (0.559)	-0.006 (0.003)	0.036 (0.064)	0.006 (0.014)	-0.029 (0.215)
Childcare × Low-Edu.	0.134 (0.042)	0.692 (0.244)	-1.570 (0.068)	-0.780 (0.291)	0.094 (0.084)	0.614 (0.387)	-0.005 (0.012)	-0.149 (0.054)	-0.016 (0.046)	0.577 (0.170)
Childcare × NBW Boy	0.020 (0.019)	0.049 (0.067)	-0.014 (0.024)	-0.170 (0.069)	0.066 (0.035)	0.013 (0.174)	-0.004 (0.005)	0.017 (0.016)	0.001 (0.016)	0.086 (0.059)
Childcare × LBW Boy	0.038 (0.035)	0.119 (0.202)	-0.079 (0.092)	-0.086 (0.145)	0.149 (0.175)	1.016 (0.579)	-0.014 (0.013)	0.026 (0.058)	0.042 (0.044)	0.111 (0.156)
Childcare × LBW Girl	0.042 (0.044)	0.111 (0.225)	-0.009 (0.064)	-0.287 (0.235)	0.141 (0.080)	0.387 (0.386)	-0.019 (0.013)	0.009 (0.066)	0.050 (0.046)	0.243 (0.228)
Test Statistics										
Weak Instrument		14.799		15.246		15.057		15.407		15.146
Wu-Hausman		1.604		6.369		2.581		1.620		2.670

Note: Parameter estimates for Equation (1). See note for Table 4.