Discussion Paper Series A No.658

Effects of Subsidized Childcare on Mothers' Labor Supply Under a Rationing Mechanism

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)

April, 2017

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

Effects of Subsidized Childcare on Mothers' Labor Supply Under a Rationing Mechanism*

Shintaro Yamaguchi† Yukiko Asai‡ Ryo Kambayashi§

April 17, 2017

Abstract

We estimate the marginal treatment effect of childcare use on mothers' labor market outcomes by exploiting a staggered childcare expansion across regions in Japan. The estimates show that the treatment effect is negatively associated with propensity to use childcare, which implies that mothers who increase their labor supply more are less likely to use childcare. Negative selection into treatment arises, because the childcare rationing rule gives preferential treatment to mothers working full-time before childcare application. These mothers are strongly attached to the labor market and likely to work regardless of the availability of subsidized childcare.

*The authors would like to thank 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. Financial supports from the Murata Science Foundation and JSPS KAKENHI Grant Numbers 15K17071, 16H02020, 16K21743 are gratefully acknowledged. The data in the Longitudinal Survey of Newborns in the 21st Century were generously provided by the Japanese Ministry of Health, Labour and Welfare.

†Department of Economics, McMaster University, 1280 Main St. W., Hamilton, ON. Canada L8S 4M4. Email: yamtaro@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

Many developed countries provide subsidized childcare to young families. One of the objectives of this policy is to increase mothers' labor supply by resolving the conflict between raising a family and pursuing a career. However, the existing evidence for the effectiveness of childcare reforms on mothers' labor supply is mixed across time and countries. Indeed, many papers find that the effect of childcare on maternal labor supply is small and/or statistically insignificant (see Lundin, Mörk, and Öckert (2008), Cascio (2009), Goux and Maurin (2010), Fitzpatrick (2010, 2012), Havnes and Mogstad (2011) and Asai, Kambayashi, and Yamaguchi (2015)).

One of the explanations for the insignificant effects of childcare reform is crowding out. Namely, subsidized formal childcare substitutes for informal childcare arrangements such as care by grandparents, and hence, mothers' labor supply does not increase. The availability and affordability of informal childcare arrangements vary across families, which means the treatment effect of childcare use is heterogeneous across families. When the supply of childcare is not large enough to take care of all children in the country, the rationing rule for childcare determines the subpopulation treated by a childcare reform. Some countries allocate childcare randomly by lottery, but other countries prioritize certain families by child's age, household income, parents' occupation, etc. Depending on the rationing rule, a childcare place may not be given to mothers who would increase their labor supply. If this is the case, the average effect of a childcare reform may be small, even though there is a subpopulation that would be strongly affected by childcare availability. This may explain why many previous papers find no effect of childcare reforms.

The objective of this paper is to estimate the heterogeneous treatment effects of childcare on mothers' labor market outcomes including participation, hours of work, earnings, and job type. We allow treatment effects to vary by unobserved propensity for childcare use by applying the marginal treatment effect (MTE) framework developed by Björklund and Moffitt (1987) and Heckman and Vytlacil (2005). Unlike the standard instrumental variable (IV) regression adopted by previous papers in the literature, the MTE framework enables us to determine which mothers are likely to change their labor supply the most and how likely they are to use childcare services, which is useful for designing an effective childcare policy.

We identify the causal effects by analyzing the childcare reform that occurred in the early 2000s in Japan. The national government legislated policies to support young families, including the expansion of subsidized childcare, in order to increase female labor supply and the fertility rate. While the national government committed to increase the supply of childcare across the country, local governments are responsible for the implementation of the policy. Because local governments differ in their policy priorities, financial status, local institutions, etc., the timing of the program rollout varies by region, which is used for identification. Because our estimation method controls for time-constant differences across regions and nationwide changes in economic conditions and policies, this identification strategy is similar to the difference-in-differences approach.

Throughout the period of analysis (and to the present day), the demand for subsidized childcare is greater than the supply in most large cities. If excess demand exists, local governments assess each family's need for childcare. While single-parent families and families on welfare receive preferential treatment, most families are ranked according to parents' working hours at the time of childcare application. This rationing rule favors full-time workers over part-time workers and job seekers. The rationale is that parents working longer hours are more in need of childcare, but the rationing rule may undermine the efficacy of the childcare reform. This is because parents working full-time at the time of childcare application are likely to use informal childcare and to work even if a subsidized place is not provided, which implies that the rationing rule is likely to cause crowding out.

Our estimates indicate that the MTEs of childcare use on market participation, hours of work, and earnings are positive for most mothers, but they are significantly heterogeneous across mothers. We find that the MTE is inversely related to the unobserved propensity for childcare use. That is, mothers with strong treatment effects are less likely to use childcare, while mothers with weak treatment effects are more likely to use childcare. Given the rationing rule that favors full-time workers, the unobserved propensity for childcare use is likely to represent an unobserved preference for work. We consider that mothers with a strong labor force attachment are willing to exert extra effort to find an informal childcare arrangement that allows them to work and raise children, even if a subsidized place is not provided, which implies that their treatment effect is weak. By contrast, mothers with weak labor force attachment may be unable to work without a subsidized place, which implies that their treatment effect is strong. We also find that this main result is robust to a number of issues including endogenous fertility, selective migration to the region in which childcare is more available, and functional form assumptions.

We then examine the consequences of the negative association between the treatment effect and the propensity for treatment in two ways. First, we calculate the average treatment effect on the treated (TT) and the average treatment effect on the untreated (TUT) by taking weighted averages of the MTE. The result indicates that TT is greater than TUT, which implies that the policy effects may improve by changing the rationing rule so that the government provides a childcare place to mothers who do not have access to one currently. Second, we evaluate the effect of a further expansion in the childcare program by counterfactual simulations under the current rationing rule. The results indicate that the policy effects become increasingly stronger as the childcare program expansion occurs. This is because those with strong treatment effects use childcare at a later stage of the childcare expansion. Overall, our analysis suggests that the current rationing rule favors mothers with stronger labor market attachment, and hence, the policy effect is undermined by crowding out.

There are two more findings worth mentioning. First, the positive effects on participation, hours of work, and earnings are brought about mainly by increasing regular employment, while nonregular employment and self-employment are not affected significantly. This result implies that not only the amount of work but also the job quality is raised by childcare enrollment. Second, the treatment effect is strongest for mothers of an infant and decreases with the child's age. This seems reasonable because informal childcare arrangements and other options such as kindergarten are available for older children.

The rest of the paper is structured as follows. Section 2 reviews the related literature. Section 3 describes the institutional background. Section 4 explains the MTE framework. Section 5 outlines the data structure and shows summary statistics. Section 6 presents our estimates of the treatment effects. Section 7 discusses the robustness of the results and shows counterfactual policy simulations. We conclude in Section 8.

2 Related Literature

Recent studies on childcare and maternal labor supply attempt to identify the causal effects using plausible exogenous variations. One of the first such papers is that by Gelbach (2002), who estimates the causal effects using the quarter of birth of five-year-old children as a source of exogenous variation. Using the 1980 US Census, Gelbach (2002) identifies the effects of the eligibility for kindergarten on maternal employment by comparing those who are barely eligible and those who are not because they were born shortly after the cutoff date to be eligible for kindergarten. He finds that the enrollment for kindergarten increased mothers' labor supply significantly. Fitzpatrick (2010) estimates the effect of childcare subsidies by applying regression discontinuity design techniques similar to those used in Gelbach (2002) to a newer cohort using the 2000 US Census and finds no subsidy effects except for single mothers. Fitzpatrick (2010, 2012) argues that childcare subsidies became less effective because the labor supply elasticity of US women has declined from 1980 to 2000.

The evidence is mixed not only across time but also across countries. Evidence from Argentina (Berlinski and Galiani (2007)), Quebec (Lefebvre and Merrigan (2008)), Spain (Nollenberger and Rodriguez-Planas (2015)), and Germany (Bauernschuster and Schlotter (2015)) shows that childcare reforms increased maternal labor supply, while evidence from Sweden (Lundin et al. (2008)), France (Goux and Maurin (2010)), Norway (Havnes and Mogstad (2011)) shows that the effect is negligibly small. Although the estimates are not directly comparable, the effect of childcare tends to be small in countries where female labor supply was already high prior to childcare reform. In those countries, the provision of formal childcare only crowds out informal childcare arrangements without affecting maternal labor supply. By contrast, in countries where the female labor force participation rate was low, a childcare reform tends to increase maternal labor supply.

The effect of childcare also varies within countries. Even in countries where the average policy effect was zero, childcare reform increased the labor supply of single mothers (see Cascio (2009), Fitzpatrick (2010), and Goux and Maurin (2010), for example). Andresen and Havnes (2016) find that Norwegian childcare reform from 2002 increased the labor supply of mothers of children aged 0–2 years, which is a younger age group than that studied in previous papers. Because many single mothers cannot afford to use other childcare arrangements and childcare for toddlers is less available than that for older children, the provision of subsidized childcare can increase the labor supply of single mothers and mothers of toddlers.

Using aggregate data at the province level in Japan from 1990 to 2010, Asai et al. (2015) estimate the intention-to-treat (ITT) effect on mothers' employment and find it to be insignificant. The current paper differs from this previous study in three important ways. First, the current paper estimates heterogeneous treatment effects using the MTE framework, while the previous paper estimates the average ITT effect only. Second, the current paper estimates the effect of childcare use by child's age, while the previous paper estimates the effect averaged over children aged 0–5 years. Given the findings by Andresen and Havnes (2016), a stronger effect is expected for younger children. Third, the current paper uses microdata after 2002, which is more recent data than that used in the previous paper. Three-generation households are common in Japan, and grandparents in the same household often take care of children while a young mother works. However, the share of three-generation households decreased from 29 percent in 1990 to 13 percent in 2010. This implies that informal childcare by grandparents has become less available, and hence, childcare is expected to have a stronger effect in more recent years.[1]

[1] Asai, Kambayashi, and Yamaguchi (2016) is a follow-up paper of their 2015 paper and estimate the ITT effect on maternal employment for the periods of 1990-2000 and 2000-2010 separately. They find insignificant effects for the 1990-2000 period and a small, but positive significant effect for the 2000-2010 period. Very recently, Nishitateno and Shikata (2017) obtain a very similar estimate for the 2000-2010 period using municipality-level data, instead of province-level data.

3 Institutional Background

3.1 Center-Based Childcare

Childcare centers in Japan are strictly regulated for quality control and are heavily subsidized. According to the Comprehensive Survey of Living Conditions, 94% of childcare centers satisfy the national quality standard set by the Child Welfare Act and are accredited by the governor of the province. Accredited childcare centers are subsidized by municipal, provincial, and national governments so that average users pay about 40% of the cost.[2] The average monthly fee per child is low at about 28,408 Yen ($\approx$ 284 USD), although it depends on age, region, household income, and the number of siblings.[3]

The remaining 6% are nonaccredited childcare centers. Some of them are owned by large employers and/or are accredited by, and receive subsidies from, local governments but not from the national government.[4] Because the vast majority of childcare centers are accredited and our main data set LSN21 does not distinguish them, we sometimes refer to accredited childcare centers as "childcare centers" for shorthand in the following.

3.2 Rationing Rule

To be eligible to use accredited childcare, parents and other extended family members under the age of 65 years and living in the same household must be unable to undertake childcare. The legitimate reasons for using an accredited childcare center include working during the day, childbearing, disability, caregiving for sick people or people with disabilities, schooling, and job search. In practice, 94.2% of parents using childcare centers satisfy the eligibility condition because they work during the day.[5]

[2] Parents of children aged 0 pay 20% of the cost, those of children aged 1 pay 30%, and those of older children pay 60%. See page 26 of Ministry of Health, Labour and Wealth (2008).

[3] See Table 7 of Ministry of Health, Labour and Wealth (2009).

[4] The fee for a child aged 0 is 46,330 JPY (about 460 USD), and that for child aged 5 is 34,161 JPY (about 340 USD). See Table 12 on page 14 of Ministry of Health, Labour and Wealth (2009).

[5] See Table 3 of Ministry of Health, Labour and Wealth (2009).

Because accredited childcare is heavily subsidized and of high quality, many parents would like to use it. When there are more applications than available places in a childcare center, the municipal government assesses the necessity of childcare use and ranks applications following the guidelines set by the Ministry of Health, Labour, and Welfare. Although details may vary across local governments, the rationing rule is largely uniform across the country.[6]

To be concrete, we describe the rationing rule for childcare using an example from the city of Yokohama. Yokohama is the largest city in Japan, with a population of 3.8 million. As is the case in major cities, there is excess demand for subsidized childcare in Yokohama. The municipality ranks applications for childcare from A (highest) to G (lowest) in the first round. Fathers and mothers are individually ranked, and the lower rank is applied to the family.

Table 1 summarizes how applications are ranked in the first round. Most parents use childcare because they usually work during the day. If a parent works outside of the home for 20 or more days per month and 8 or more hours per day, he/she is ranked highest at A. The rank is lowered if a parent works less, and a rank of C is given if he/she works 16 or more days per month but works 4–7 hours per day. A parent who is currently not working but has a job offer is ranked lower than those currently working, and a rank of D is given if he/she works for 16 or more days per month and 7 hours per day. A parent working for fewer than 16 days per month or 4 hours per day is not considered to have a legitimate reason to use childcare. The days and hours of work are assessed 1 month before the childcare application and must be verified by a document signed by the employer. Those who are on parental leave can report their days and hours of work before they took leave.

Although a parent can use accredited childcare if he/she is searching for a job, he/she is given the lowest rank of G. This is problematic for some job seekers. Some employers may be unwilling to hire a young mother who does not find a childcare arrangement, because she may not be able to work without childcare. Employment is required for childcare use, but childcare use may also be essential for finding

[6] We confirm that the rationing rules are very similar among the largest cities, where rationing typically takes place. These cities include Yokohama, Osaka, Nagoya, Sapporo, Kobe, Kyoto, Fukuoka, Kawasaki, Saitama, Hiroshima, and Sendai.

employment.

Lone parents and parents with disabilities are given the highest priority, and hence, a rank of A. The rank is raised by one unit if the family is on welfare, the main earner lost his/her job, etc.; however, household income has little influence on childcare allocation otherwise.

If many applications are ranked equally at the cutoff level, the municipal government further ranks these applications in the second and third rounds. For example, if older siblings are already enrolled in the same childcare center, extra points are given to the application in the second round.

Table 1: Necessity Assessment in the First Round (Yokohama, 2010)

Reason	Note	Rank
Work	$\geq$ 20 days/month and $\geq$ 8 hours/day	A
	$\geq$ 16 days/month and $\geq$ 7 hours/day	B
	$\geq$ 16 days/month and 4-7 hours/day	C
	$\geq$ 16 days/month and $\geq$ 7 hours/day (Job Offer)	D
	$\geq$ 16 days/month and 4-7 hours/day (Job Offer)	E
	(Rank is lowered by one unit if work at/from home.)	
Job Search	Up to 3 months	G
Single Parent	If engaged with work, training, or job search	A
Disability	Class 1 or 2	A
Childbearing	8 weeks before and after	D
School		D
⋮		⋮

Source: Page 12 of Aoba Ward, City of Yokohama (2009).
Note: If the family is on welfare, the rank is raised by one unit. If many households are in the same rank at the threshold level for childcare allocation, the application is further ranked for tie-break in the second and third rounds.

3.3 Childcare Expansion

The national government recognized the shortage of accredited childcare centers in the early 1990s. It launched the Angel Plan (1994–1998) and the New Angel Plan (1999–2003), which included an expansion of childcare capacity, extension of childcare service hours to include weekends and holidays, and subsidies to promote the take-up of parental leave and shorter working hours. Unfortunately, these two plans were too small and failed to increase the supply of accredited childcare. In 2003, the national government enacted the Basic Act for Measures to Cope with Society with Declining Birthrate and committed to taking legal and financial measures to increase the supply of childcare. This policy reform increased the number of accredited childcare places by 12% from 2000 to 2010.

While the national government covers half of the cost, the provincial and municipal governments are responsible for the rollout of the childcare reform. Depending on the financial status and policy priorities, the pace of the program rollout varied by region. This variation in the pace of childcare expansion across regions is exploited for the identification of causal effects. We will discuss this identification strategy extensively in Section 4.2 after we describe the MTE framework.

4 Analytical Framework and Empirical Strategy

4.1 Marginal Treatment Effect

4.1.1 Setup

Using the MTE framework developed by Björklund and Moffitt (1987) and Heckman and Vytlacil (2005),[7] we estimate heterogeneous treatment effects varying by observed and unobserved characteristics of mothers.

Define $j \in \{0,1\}$ as an index for treatment status such that $j=1$ implies being treated and $j=0$ implies being untreated. A potential outcome Y_j for treatment status j is given by

$$Y_j = X\beta_j + U_j \tag{1}$$

$$E(U_j|X) = 0, \tag{2}$$

where X is a vector of control variables. Following Brinch, Mogstad, and Wiswall (forthcoming), we assume a zero conditional mean of errors U_j instead of indepen-

[7] Cornelissen, Dustmann, Raute, and Schönberg (2016) is an excellent introduction to the MTE framework for applied researchers.

dence.

Treatment status is determined by the following selection equation

$$D = 1\{X\gamma + \delta Z - V > 0\}, \tag{3}$$

where D is a dummy variable for treatment, $1\{\cdot\}$ 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 is the childcare coverage rate, which is defined as childcare slots per child in a given region. The validity of the instrument is discussed extensively in Section 4.2. We also include the interactions of the coverage rate and a subset of exogenous variables X in the instruments. Because a larger value of V keeps more mothers from treatment, we refer to it as a resistance to treatment.

The selection equation (3) can be rewritten as

$$\begin{aligned} D &= 1\{X\gamma + \delta Z > V\} && (4)\\ &= 1\{F_V(X\gamma + \delta Z) > F_V(V)\} && (5)\\ &= 1\{P(X\gamma + \delta Z) > U_D\}, && (6) \end{aligned}$$

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

The MTE is defined as

$$MTE(X = x,\ U_D = u_D) = E(Y_1 - Y_0 | X = x,\ U_D = u_D). \tag{7}$$

It is interpreted as the gain from treatment for a mother whose observed characteristics are $X = x$ and the quantile of the unobserved resistance to treatment is $U_D = u$. Policy relevant parameters such as the average treatment effect (ATE), the treatment effect on the treated (TT), the treatment effect on the untreated (TUT), and the local average treatment can be derived as weighted averages of the MTE.

What is the economic interpretation of u_D in the context of this research? Recall how the local governments select successful applications. The rationing rule sorts families by how much parents work. Because most fathers in the sample work full-time, mothers' working hours are the key determinant of the probability of childcare use. Mothers' unobserved preference for work and their skills are likely to be the main components of the unobserved resistance for work. Specifically, a low u_D implies strong labor market attachment, while a high u_D implies weak labor market attachment.

4.1.2 Local Instrumental Variable Estimator

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

$$MTE(X=x,\ U_D=u_D) \quad = \quad x(\beta_1-\beta_0)+E(U_1-U_0|U_D=u_D). \tag{8}$$

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

$$E(Y|X=x,\ P(X,Z)=p) \quad = \quad x\beta_0+x(\beta_1-\beta_0)p+K(p), \tag{9}$$

where $K(p)$ is a nonlinear function of the propensity score. The MTE for the mother with $X=x$ and $U_D=p$ is given by the derivative of Equation (9) with respect to the propensity score,

$$\begin{aligned} MTE(X=x,\ U_D=p) \quad &= \quad \frac{\partial E(Y|X=x,\ P(X,Z)=p)}{\partial p} && (10)\\ &= \quad x(\beta_1-\beta_0)+\frac{\partial K(p)}{\partial p}. && (11) \end{aligned}$$

How does the local IV estimator identify the MTE defined by unobserved characteristics u_D? We ignore observed characteristics X to simplify the argument for the moment. The selection equation (6) implies that those with the unobserved characteristics $u_D<p$ are selected into treatment and that 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. Note that the average outcome as expressed in Equation (9) changes in response to the change in the propensity score p. The MTE is given by the change in the average outcome divided by the fraction of individuals newly selected into treatment.

4.2 Threat to Identification

Our instrument is the childcare coverage rate, which is defined as childcare places per child in a given region. Because our control variables include year and region dummies, time-invariant regional characteristics and nationwide changes in the social and economic conditions are accounted for. The variation exploited for identification is the growth of the childcare coverage rate in a given region, which is equivalent to the difference-in-differences approach.

One of the identifying assumptions is that the growth of the coverage rate is exogenous. To see whether this assumption is valid, we examine the potential determinants of the growth of the coverage rate. According to Cabinet Office (2010), there are three obstacles preventing childcare expansion. First, the regulations are strict and uniform across the country, even though some (e.g., area per child and requirement for a kitchen) are unnecessary or unrealistic for urban areas. Second, the local governments do not have a permanent budget for childcare operations, although the national government transfers some additional funds to the local governments temporarily. Third, some municipal governments cannot acquire the land to build new childcare centers, because the rents and land prices are prohibitively high in urban areas.

We assess how these factors and other regional characteristics affect the pace of childcare expansion. We regress the change of the coverage rate from 2000 to 2010 on several prereform regional characteristics including 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.

Table 2 reports the regression results. The female labor force participation rate in the prereform period is positively and significantly correlated with the growth of the childcare coverage rate. The demand for childcare is considered to be high in

Table 2: Determinants of the Growth of Childcare Coverage Rate

	Change in Coverage Rate
Female Labor Force Participation Rate	0.186
	(0.089)
Total Fertility Rate	−0.001
	(0.050)
Financial Capability Index	−0.062
	(0.035)
Log Land Price	0.009
	(0.015)
Log Average Female Wage	−0.060
	(0.099)
Num. obs.	80

Sources: All explanatory variables are measured in 2000. 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, I omitted City of Yokosuka and non-major cities in Province of Kanagawa, although they are included in the main analysis.

the region where the female labor force participation is high. The estimate suggests that the supply of accredited childcare increased to meet the high demand in such a region. The total fertility rate in the prereform period has no effect on the growth of the coverage rate. Despite the argument by the Cabinet Office (2010), the financial status of the provincial government, the land price, and the female wage rate in the prereform period do not affect the growth of the coverage rate significantly.

This result indicates that the growth of the coverage rate is not completely random, and hence, accounting for potential policy endogeneity is necessary for unbiased estimates. To address this issue, we include interactions of the prereform regional characteristics and year dummies in our control variables.

Another threat to identification is selective migration. A popular narrative says that obtaining a place in an accredited childcare center is extremely difficult in Tokyo and that some people even move to other districts for childcare. Using the Employment Structure Survey[8] 2012, we take a sample of mothers of children under age 6 and examine their reasons for the most recent move and where they moved from. We find that "for childrearing and education," 9.5% moved within the same city, 4.6% moved from another city in the same province, and 1.4% moved from another province. Because a region in this study is defined by the combination of city and province, at most 4.6% moved between regions for childcare. As we will show in Section 7.1, selective migration seems to have little effect on the estimates.

There was another major policy reform that may have affected female labor supply during the period of analysis. In 2005, nonregular workers became eligible for a paid one-year parental leave. In addition, the replacement rate of cash benefits was raised gradually from 25% to 50% from 2000 to 2007.[9] We do not believe that this policy reform biases our estimates, because it is uniformly legislated across the country and its effect on mothers' labor supply is accounted for by the year dummy. We are not aware of any major policy changes that may affect female labor supply at the region level.

Other issues that might affect our estimates include endogenous fertility, presence of siblings, and functional form assumptions. These issues are discussed ex-

[8]It is conducted by the Statistics Bureau every three years and covers about 1% of the population.
[9]See Asai (2015) and Yamaguchi (2016) for evaluation of these policy changes.

tensively in Section 7.1, but our main results are largely unaffected.

4.3 Empirical Implementation

We implement the local IV estimator as follows. In the first step, we estimate the propensity score using a flexible probit model. The covariates include the coverage rate up to the third-order term, parents' age and education, and year and region dummies. To allow for heterogeneous responses to the coverage rate, we interact the coverage rate and parents' characteristics. In addition, we include the interactions of year dummies and prereform regional characteristics to address policy endogeneity.

In the second step, we estimate the outcomes equation (9) using a linear regression with the assumption that $K(p)$ is a quadratic function. We allow for higher-order terms of the propensity score in the robustness checks. To allow for heterogeneous treatment effects by parents' age and education, we interact them with the propensity score. Standard errors are bootstrapped with 100 replications and clustered at the region level. Note that one replication in our bootstrap procedure includes both the first and second steps so that uncertainty in the estimates in the first step is taken into account.

5 Data

5.1 Data Sources

Our main data source is LSN21, which is a census of children who were born January 10–17, 2001, July 10–17, 2001, and May 10–24, 2010. The first survey is conducted when the child is 6 months old. Subsequent questionnaires are completed every year about 6 months after their birthdays. The response rates are high at 93.5 and 88.1 percent in the first survey years for cohorts born in 2001 and 2010, respectively. About 83 percent of respondents in the first survey remain in the survey 3 years later at age 3.5. These response rates for the LSN21 are higher than those for the National Longitudinal Survey of Children and Youth,[10] which

[10] In the National Longitudinal Survey of Children and Youth in Canada, the response rate in the first cycle conducted in 1994/1995 was 86.5%, and 67.8% of the children in the original cohort

provides Canadian longitudinal data used by Baker, Gruber, and Milligan (2008), among others.

We draw data on accredited childcare centers from the annual Report on Social Welfare Administration and Services, which covers all provinces and major cities with a population of over 200,000. We define region by a combination of city and province: a region is a major city or a set of all municipalities in a province except for the major cities. We include 82 regions that are included in the data in both 2002 and 2011, which consists of 45 provinces and 37 major cities. The provinces of Fukushima and Miyagi were omitted because of missing data, as they were severely affected by the Great East Japan Earthquakes and Tsunami in 2011.

Child population is taken from the quinquennial Census. For the years when the Census is not conducted, we estimate child population by linear interpolation. Other regional characteristics in 2000 are drawn from various sources. See the note on Table 2 for details.

5.2 Descriptive Statistics

Table 3 shows the summary statistics of the sample. The childcare enrollment rate is low at 3.6% when the child is 0.5 years old, but it increases to 23.4% 1 year later. As children grow older, the enrollment rate increases to 36.8% when the child is 3.5 years old.

The coverage rate, defined as the number of places per child, is about 0.3 and increases gradually with the child's age, which reflects the progress of the childcare reform. The average coverage rate is higher for the treated than for the untreated, implying that the coverage rate is positively associated with the childcare enrollment rate.

Parents' ages are evaluated when the child is 0.5 years old. The average age of mothers is 30.405, and fathers are about 2 years older on average than mothers. Parents' education is measured when the child is 1.5 years old. About 5% of parents are less-than-high-school educated and about one-third of parents are high school educated. Two-year college is the most common education level for mothers, and

responded in cycle 3 conducted in 1998/1999.

about 20% of mothers went to university. In contrast, four-year university is the most common education level for fathers, and about 17% of fathers went to two-year college. We do not find large systematic differences in mothers' characteristics between the treated and untreated when other characteristics are not controlled for. In contrast, father's characteristics are substantially different between the treated and untreated. The treated fathers are younger and less educated than the untreated fathers.

The fraction of working mothers is 13.3% when the child is 0.5 years old. Note that many mothers are entitled to 1 year of paid maternity leave until their child reaches the age of 1 year. Many mothers return to the labor market after their paid leave period, and 33.7% of mothers work when the child is 1.5 years old. The fraction of working mothers increases with child's age, and 42.8% of mothers work when the child is 3.5 years old. There is a strong association between the use of center-based childcare and mothers' labor market participation. The labor market participation rate of childcare users is high at 83–91%, but that of nonusers is only 10–20%.

Note that the fraction of working mothers is higher than the enrollment rate at the same age, implying that many mothers work without using formal childcare. The labor market participation rate among mothers not using childcare increases with child's age, suggesting more availability of informal childcare for older children.

The mean number of working hours increases with child's age, although they are not observed at age 1.5 years. The mean number of working hours among childcare users is stable at about 30 hours per week, while that among nonusers is about 2–5 hours per week. Mean earnings are not monotonically increasing with child's age. The mean earnings of childcare users are higher than those of nonusers. Earnings are not observed at age 2.5 years.

Market work is categorized into either regular work, nonregular work, or self-employment. Regular employment is typically under a permanent contract and a full-time job, while nonregular employment is typically under a limited-term contract and a part-time job. They also differ in hourly wages, nonwage benefits, employer-sponsored training, and eligibility for the mandated PL (see Kambayashi

Table 3: Summary Statistics

	All			Comparison by Treatment		
	Nobs.	Mean	S.D.	Treated	Untreated	p-value for Difference
Childcare Enrollment						
Age 0.5	72475	0.036	0.188	1.000	0.000	
Age 1.5	72477	0.234	0.423	1.000	0.000	
Age 2.5	67929	0.314	0.464	1.000	0.000	
Age 3.5	65242	0.368	0.482	1.000	0.000	
Coverage Rate						
Age 0.5	72475	0.305	0.111	0.328	0.304	0.000
Age 1.5	72477	0.313	0.113	0.336	0.306	0.000
Age 2.5	67929	0.321	0.116	0.348	0.309	0.000
Age 3.5	65242	0.331	0.118	0.360	0.313	0.000
Mother's Characteristics						
Age	72475	30.405	4.516	30.663	30.395	0.003
Less Than High School	72475	0.046	0.210	0.057	0.046	0.012
High School	72475	0.334	0.472	0.343	0.334	0.314
2-Yr College	72475	0.419	0.493	0.431	0.419	0.233
4-Yr University or Higher	72475	0.200	0.400	0.169	0.201	0.000
Father's Characteristics						
Age	72475	32.402	5.524	32.565	32.396	0.146
Less Than High School	72475	0.073	0.261	0.109	0.072	0.000
High School	72475	0.354	0.478	0.439	0.350	0.000
2-Yr College	72475	0.169	0.375	0.176	0.168	0.307
4-Yr University or Higher	72475	0.404	0.491	0.276	0.409	0.000
Market Work						
Age 0.5	71817	0.133	0.339	0.878	0.104	0.000
Age 1.5	72046	0.337	0.473	0.912	0.161	0.000
Age 2.5	67338	0.374	0.484	0.873	0.146	0.000
Age 3.5	64215	0.428	0.495	0.828	0.195	0.000
Hours Per Week Worked						
Age 0.5	71957	3.255	10.717	30.921	2.215	0.000
Age 2.5	67338	12.503	18.704	31.317	3.900	0.000
Age 3.5	64010	13.587	18.742	28.884	4.761	0.000
Earnings (mil. JPY)						
Age 0.5	70552	0.741	1.557	2.135	0.691	0.000
Age 1.5	67555	0.451	1.098	1.458	0.183	0.000
Age 3.5	61746	0.797	1.561	1.802	0.245	0.000
Regular Work						
Age 0.5	71817	0.041	0.199	0.493	0.024	0.000
Age 1.5	72046	0.168	0.374	0.541	0.054	0.000
Non-Regular Work						
Age 0.5	71817	0.039	0.193	0.303	0.029	0.000
Age 1.5	72046	0.113	0.316	0.305	0.054	0.000
Self-Employed						
Age 0.5	71817	0.040	0.195	0.069	0.039	0.000
Age 1.5	72046	0.042	0.200	0.058	0.037	0.000

Source: LSN21.
Note: The sample includes two-parent families. Parents' ages and education are measured when child is 0.5 and 1.5 year old, respectively. Not all labor market outcomes are available at all ages.

and Kato (2013)). Information on employment type is available when the child is aged 0.5 and 1.5 years. The fraction of regular workers is slightly higher than that of nonregular workers, and only 4% of mothers are self-employed.

Table 4: Share of Childcare Mode for Working Mothers

Care Mode	Childcare Center	Grandparents	Sitter/ Nannies	Parents Only
Age 0.5	0.241	0.334	0.062	0.362
Age 1.5	0.633	0.198	0.031	0.138
Age 2.5	0.733	0.150	0.020	0.098
Age 3.5	0.712	0.101	0.070	0.117

Source: LSN21

Note: 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 fulltime. If a child is cared by 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 is reported, parents are the primary caregiver.

Many working mothers use informal childcare arrangements, and their choice of childcare mode changes with child's age. Table 4 shows the share of childcare mode for working mothers. When the child is 0.5 years old, only 24.1% of working mothers use a childcare center. Most children at age 0.5 years are cared for by their parents or grandparents. The use of babysitters and nannies is rare, at 6.2%.

While most working mothers of infants do not use formal childcare, it is the most common childcare mode for working mothers of older children aged 1.5–3.5 years. At age 1.5 years, 63.3% of working mothers use childcare, and more than 70% of working mothers use childcare when the child is 2.5–3.5 years old. The use of informal care by grandparents becomes less common as the child grows older. The use of babysitters or nannies is rare regardless of the child's age. About 10% of the children of working mothers are cared for by parents only.

6 Results

We show our main results in three steps so that the econometric model becomes increasingly complex. First, in Section 6.1, we show the estimates for the first-stage and reduced-form regressions. Our estimates indicate that the childcare coverage rate indeed increases childcare enrollment and maternal employment rates, which implies that our instrument, the childcare coverage rate, is not weak. Second, in Section 6.2, assuming an homogeneous treatment effect for tractability, we show the estimates for the treatment effect of childcare enrollment by conventional IV regression and bivariate probit model. We hope that the analysis here is useful for readers, because these econometric methods are used routinely. In addition, the estimates can be compared with those from previous studies that use IV regressions. Third and finally, we allow for heterogeneity in the treatment effect in Section 6.3. There, we show estimates for the marginal treatment effects that vary by unobserved propensity for childcare enrollment.

6.1 Effects of Childcare Coverage Rate on Enrollment and Maternal Employment

We begin our analysis by presenting region-level evidence. The panels in the left column of Figure 1 present the relationship between the region-level coverage rate and childcare enrollment rate by child's age, after removing the effects of prereform regional characteristics, parental characteristics, and region and year fixed effects (see the note on Figure 1 for details). The panels in the right column show the relationship between the coverage rate and mothers' employment rate at the region level.

In all graphs, the slope is positive and statistically significant at the 10% level for all ages from 0.5 to 3.5 years. Note that the results do not seem to be driven by an outlier. The graphs provide prima facie evidence that the expansion of childcare places increases the enrollment of children and the employment of their mothers.

These graphs make our identification strategy transparent; however, the residual plot can only help us visualize the linear relationship. The marginal effect of the

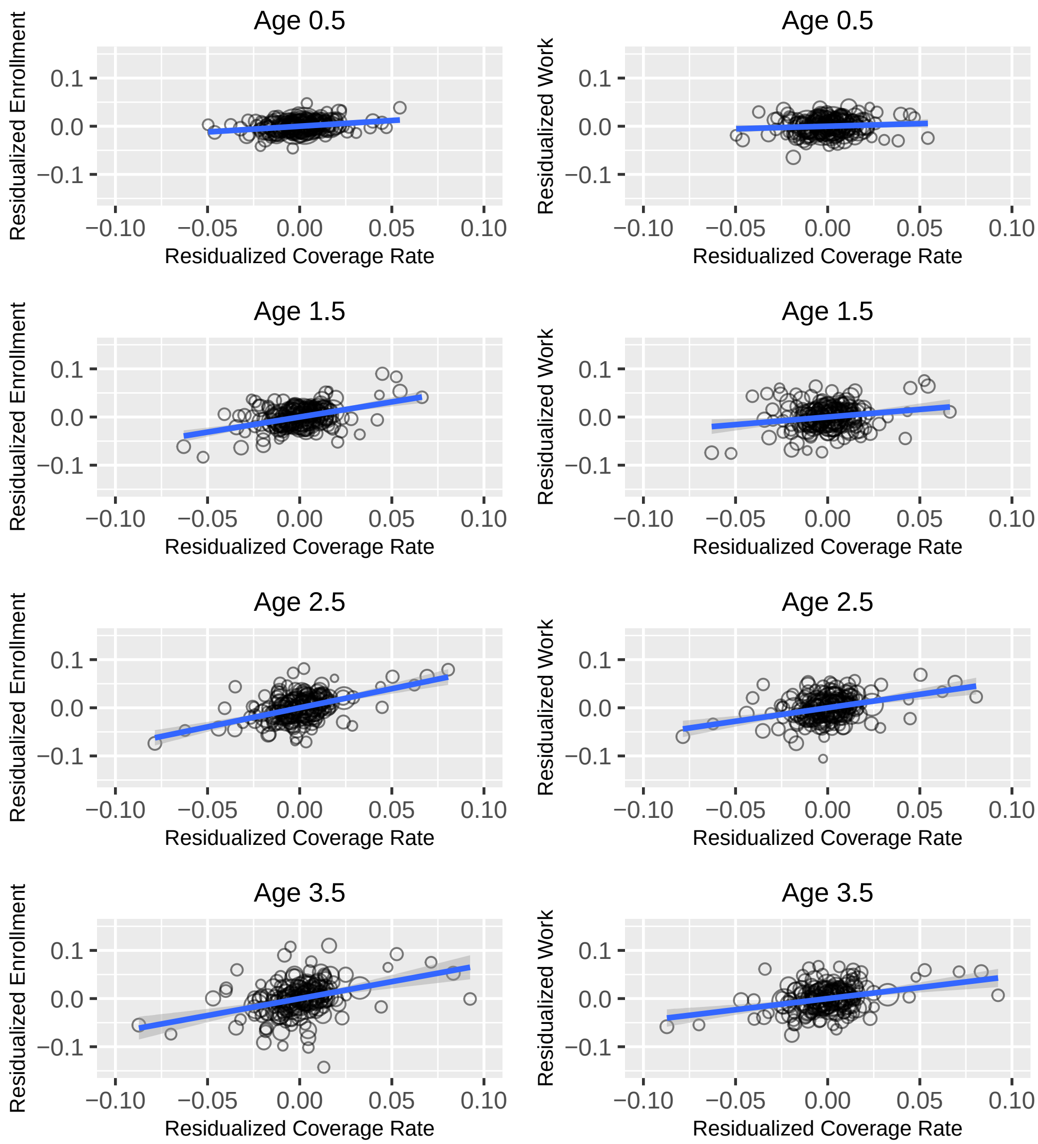

Figure 1: Coverage Rate and Childcare Enrollment by Region
Source: Authors' calculation based on LSN21.
Note: Residualized coverage, childcare enrollment, and maternal employment rates are calculated as follows. We regress each of the variables on parents' age and education, regional characteristics at the 2000, and year and region dummies, and then, take averages of the residuals by region. The size of the bubbles indicates the number of observations in the region. The fitted line is estimated by the weighted least squares and shown with the 90% confidence interval.

coverage rate on the enrollment rate may vary by the coverage rate, but that is masked in these graphs. To examine potential nonlinearity between the coverage rate and outcomes, we estimate a flexible probit model that includes up to third-order polynomials of the coverage rate and parents' and regional characteristics.

The top-left panel of Figure 2 shows the effects of the coverage rate on childcare enrollment by child's age. The enrollment rate increases with the coverage rate for all age groups, which confirms the results from the region-level analysis. In addition, the graph indicates that the enrollment rate changes with the coverage rate. The graph for children aged 3.5 years shows a concave profile, implying that their enrollment rate does not increase much at a higher coverage rate. This is presumably because many 3.5-year-old children choose to enroll in kindergarten instead of childcare, even if childcare is available. Kindergarten is not an option for younger children aged 0.5–2.5 years. By contrast to 3.5-year-old children, their enrollment increases with the coverage rate at an increasing rate.

The top-right panel of Figure 2 shows the effects of the coverage rate on the fraction of working mothers. It increases with the coverage rate and the child's age. Note that the fractions of working mothers and childcare users do not generally coincide for two reasons. First, some working mothers use informal childcare arrangements. Second, the municipal governments accept applications from mothers who do not work but are unable to care for their children because of sickness, disability, schooling, looking for work, etc. These eligible but nonworking mothers are given low priority but are able to use childcare when the coverage rate is high.

To see whether an expansion of childcare crowds out informal childcare arrangements for working mothers, we estimate probit models for the joint choice of mothers' work and childcare mode. The bottom-left panel shows how the coverage rate affects the fraction of mothers who work and use a childcare center. The fraction increases with the coverage rate and the child's age. The graph is concave for 3.5-year-old children, while it is convex for younger children. The bottom-right panel shows how the coverage rate affects the fraction of mothers who work and use an informal childcare arrangement. It decreases with the coverage rate, but no clear pattern can be seen by the child's age.

Many working mothers rely on informal childcare arrangements when the cov-

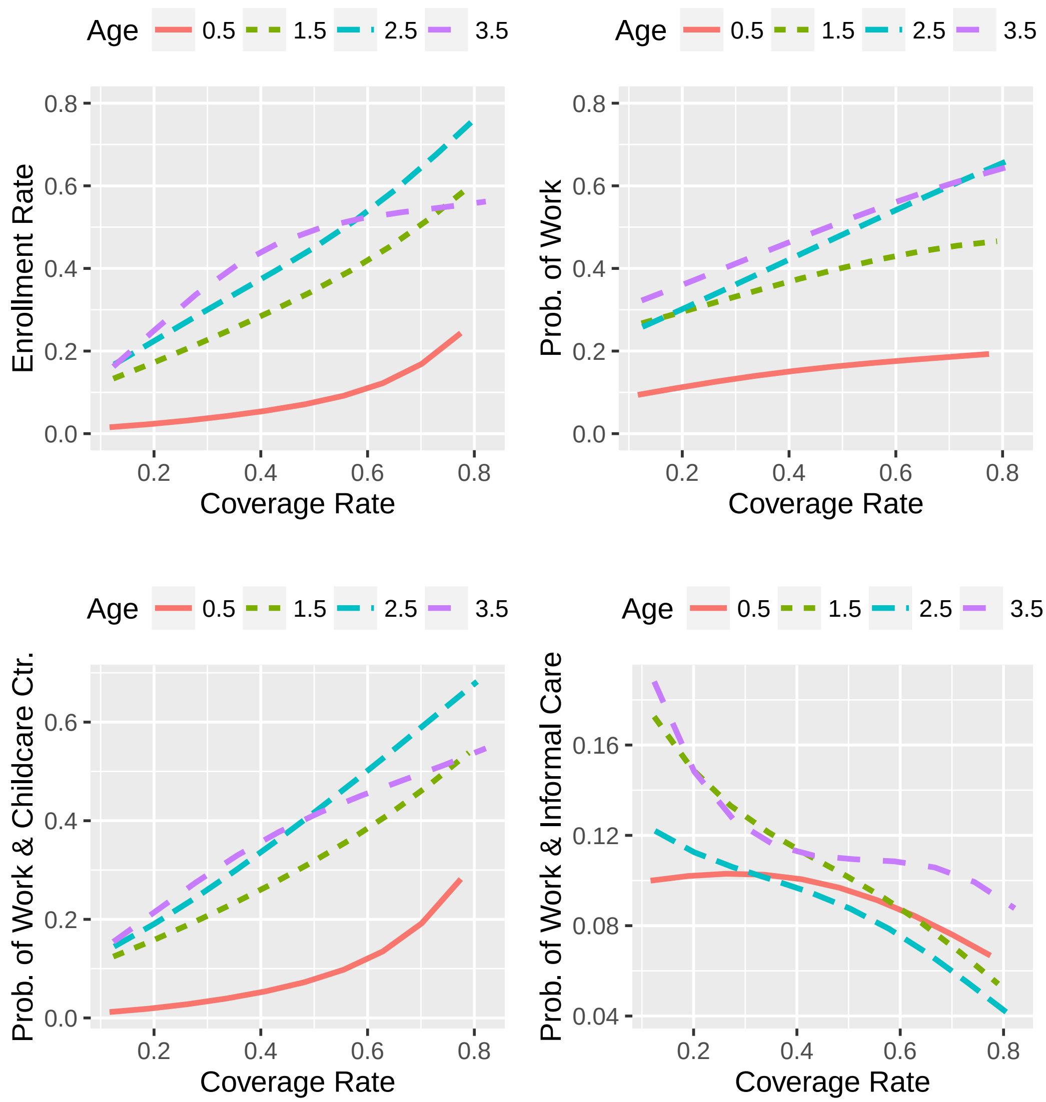

Figure 2: Coverage Rate, Childcare Enrollment, and Mother's Work

Note: Estimates are based on the probit model in which the explanatory variables include the coverage rate up to the third-order polynomial, parents' age and education, regional characteristics in 2000, and region and year fixed effects. The polynomials of the coverage rate are interacted with parents' characteristics.

erage rate is low, but informal childcare is increasingly substituted with formal childcare centers as the coverage rate increases. This result implies that the expansion of childcare supply crowds out informal childcare among working mothers. Even though parents' work is effectively a prerequisite for the use of a childcare center, providing a new childcare place does not necessarily add a mother to the labor market, because some of them simply substitute informal childcare arrangements and continue to work.

Childcare enrollment and mothers' work are also affected by parents' age and education. Table 5 shows the average marginal effects of parental characteristics based on the probit model.

Mothers' education is positively correlated with the enrollment rate. Remember that parents' work is virtually required for childcare use. Skilled mothers are more likely to work and to use childcare, because their opportunity cost of staying at home is high. In addition, the rationing rule gives a higher priority to mothers with stronger labor market attachment, which is positively correlated with education.

Father's age and education are negatively correlated. If a wife takes her husband's earnings as given, her labor supply and childcare use are negatively correlated with her husband's earnings because of the income effect, which accounts for why childcare use is negatively correlated with father's age and education.

We show that the childcare coverage rate increases the mothers' employment rate, but the interpretation of the result is not necessarily straightforward. This is because the effect of the coverage rate on mothers' employment is the product of the effect on childcare enrollment and the treatment effect of childcare enrollment. In the following subsections, we show estimates for the treatment effect of childcare enrollment.

6.2 Estimating Treatment Effects by IV Regression and Bivariate Probit

In this subsection, we present estimates for treatment effects on mothers' labor market outcomes using a conventional IV regression and bivariate probit model, assuming that treatment effects are homogeneous. Because these methods are well

Table 5: Marginal Effects of Coverage Rate on Childcare Enrollment

	Age 0.5	Age 1.5	Age 2.5	Age 3.5
Childcare Enrollment				
Mother's Age	0.001	0.002	0.001	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)
Less Than HS (Mother)	-0.001	-0.095	-0.098	-0.075
	(0.004)	(0.008)	(0.010)	(0.010)
HS (Mother)	-0.007	-0.099	-0.104	-0.078
	(0.002)	(0.005)	(0.006)	(0.006)
2-Yr College (Mother)	-0.001	-0.055	-0.071	-0.058
	(0.002)	(0.005)	(0.005)	(0.005)
Father's Age	-0.000	-0.002	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
Less Than HS (Father)	0.034	0.111	0.148	0.185
	(0.004)	(0.007)	(0.008)	(0.008)
HS (Father)	0.020	0.056	0.078	0.101
	(0.002)	(0.004)	(0.004)	(0.005)
2-Yr College (Father)	0.013	0.055	0.072	0.091
	(0.002)	(0.005)	(0.005)	(0.005)
Mother's Work				
Mother's Age	0.003	0.004	0.004	0.005
	(0.000)	(0.001)	(0.001)	(0.001)
Less Than HS (Mother)	0.011	-0.144	-0.124	-0.071
	(0.007)	(0.009)	(0.010)	(0.011)
HS (Mother)	-0.014	-0.136	-0.119	-0.078
	(0.004)	(0.006)	(0.006)	(0.006)
2-Yr College (Mother)	-0.008	-0.083	-0.072	-0.054
	(0.004)	(0.005)	(0.005)	(0.005)
Father's Age	0.002	-0.001	-0.001	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)
Less Than HS (Father)	0.111	0.160	0.183	0.222
	(0.006)	(0.008)	(0.008)	(0.008)
HS (Father)	0.059	0.093	0.111	0.127
	(0.003)	(0.004)	(0.005)	(0.005)
2-Yr College (Father)	0.059	0.091	0.098	0.115
	(0.004)	(0.005)	(0.005)	(0.006)

Note: Standard errors are in parenthesis. Estimates are based on the probit model in which the explanatory variables include the coverage rate up to the third-order polynomial, parents' age and education, regional characteristics in 2000, and region and year fixed effects. The polynomials of the coverage rate are interacted with parents' characteristics. The reference group for education is those with 4-year university education or higher.

known and commonly used, our estimates can be compared with those in the previous studies in the literature.

We use the propensity score as an instrument rather than the coverage rate itself. The propensity score is a function of the coverage rate and other exogenous variables, but the use of the propensity score as an instrument has advantages over the use of the coverage rate in our context. First, it is required by the local IV estimator for the MTE. By using the same instrument across different models (IV, probit, and local IV models), the estimates become comparable. In particular, an IV estimate can be interpreted as a weighted average of the MTE (see Heckman and Vytlacil (2005)). Second, the propensity score provides the efficient IV if it is correctly specified (see Wooldridge (2010)). Note that this method produces a consistent IV estimate even if the propensity score is not modeled correctly.

The first column of Table 6 shows the effects of childcare use on mothers' work estimated by OLS. The estimates range from 0.619 to 0.754, and the treatment effects decrease with age. Of course, they cannot be interpreted as causal because of possible endogenous selection into treatment.

The estimates from the IV regression are reported in the second column. The estimated effect on mothers of children aged 0.5 years is incredibly large. Given that only about 13% of them work and only about 3% of them use a childcare center (see Table 3), the implausible estimate is caused by poor approximation by the linear probability model. For this age group, accounting for nonlinearity by the probit model seems to be important. The IV estimates for other age groups range from 0.426 to 0.685 and are smaller than the OLS estimates. The treatment effects decrease with child's age. Note that we soundly reject the null hypothesis that our instrument is weak, as shown in the third column.

To account for nonlinearity, we estimate the treatment effects using probit models. The fourth column shows the estimates of a univariate probit model, which does not address possible endogeneity biases. They are very similar to estimates by OLS. The fifth and last column shows estimates of a bivariate probit model that is comparable to the IV regression as they both use an instrument. The bivariate-probit estimates are smaller than those of a univariate probit for all age groups. Moreover, the difference between univariate and bivariate probit estimates increases with child's

Table 6: Treatment Effects on Mother's Market Work

	Linear Models			**Nonlinear Models**	
	OLS	IV	F-stat for Weak IV	Probit	Bivariate Probit
Age 0.5	0.754	1.142	206.639	0.752	0.731
	(0.008)	(0.155)		(0.007)	(0.064)
Age 1.5	0.744	0.685	147.599	0.744	0.690
	(0.010)	(0.086)		(0.003)	(0.040)
Age 2.5	0.723	0.661	95.613	0.723	0.498
	(0.012)	(0.089)		(0.003)	(0.069)
Age 3.5	0.619	0.426	106.285	0.618	0.144
	(0.015)	(0.108)		(0.003)	(0.070)

Note: Standard errors are in parenthesis and those for OLS and IV are clustered at the region level. The dependent variable is a dummy for mother's market work. The endogenous variable is a dummy for the use of a childcare center. The IV is the propensity score estimated by the flexible probit (see Section 6.1). Other exogenous variables include parents' age and education, regional characteristics in 2000, and year and region fixed effects.

age, suggesting that endogeneity bias is more severe for older children.

Although estimates vary between the IV regression and the bivariate probit, they share some important features: the estimated treatment effect is positive and significant, and decreases with child's age. The treatment effect is smaller for older children, because informal childcare is more readily available. In particular, children aged 3.5 years are eligible to enroll in kindergarten. Providing a childcare place to an older child does not change his/her mother's work hours, because it only crowds out the existing informal childcare. Our estimates from the IV and bivariate probit models help to interpret the ITT effects. The ITT effects are small for mothers of very young children, because they are less likely to use a childcare center, even though their treatment effects are greater than those of the mothers of older children.

6.3 Estimating Marginal Treatment Effect by Local IV Regression

We estimate the MTE using the local IV estimator outlined in Section 4.1.2. Compared with the IV regression and bivariate probit models, the local IV regression is more flexible because we allow for heterogeneous treatment effects using unobserved characteristics. In the following, we discuss the results for mothers of children aged 1.5–3.5 years. The estimator does not provide reasonable estimates for mothers of children aged 0.5 years, because both the enrollment rate and the fraction of working mothers are very low, and the linear probability model in the second stage regression provides a poor approximation.

Figure 3 shows how the MTE on various labor market outcomes changes with the quantile of the unobserved resistance to treatment u_D. The MTE is averaged over observed characteristics and shown with the 90% confidence interval. The three panels in the top row show the MTE curves for mothers' market work by child's age. They are all significantly above zero and increase with the resistance to treatment, which implies a reverse selection on the treatment effect. Remember that those with a lower value of resistance to treatment are more likely to be enrolled in childcare. The upward-sloping MTE curves imply a negative relationship between the treatment effect and the propensity for childcare use; that is, those with weak treatment effects are more likely to be given a childcare place, while those with stronger treatment effects are less likely to be given a childcare place.

The panels in the second row show the MTE on weekly hours of work. Work hours are not included in the data when the child is age 1.5 years. Again, the MTE is positive and significant and increases with the unobserved resistance. The panels in the third row show the MTE on annual earnings in million JPY ($\approx$ 10,000 USD). The MTE on earnings is positive and significant except for very low values of u_D and increases with the unobserved resistance.

Figure 4 shows the MTE on the choice of employment types that are unavailable when the child is 2.5 and 3.5 years old. Market work is categorized into either regular work, nonregular work, or self-employment. A typical regular worker is a full-time worker, better paid, with more fringe benefits, and under a permanent con-

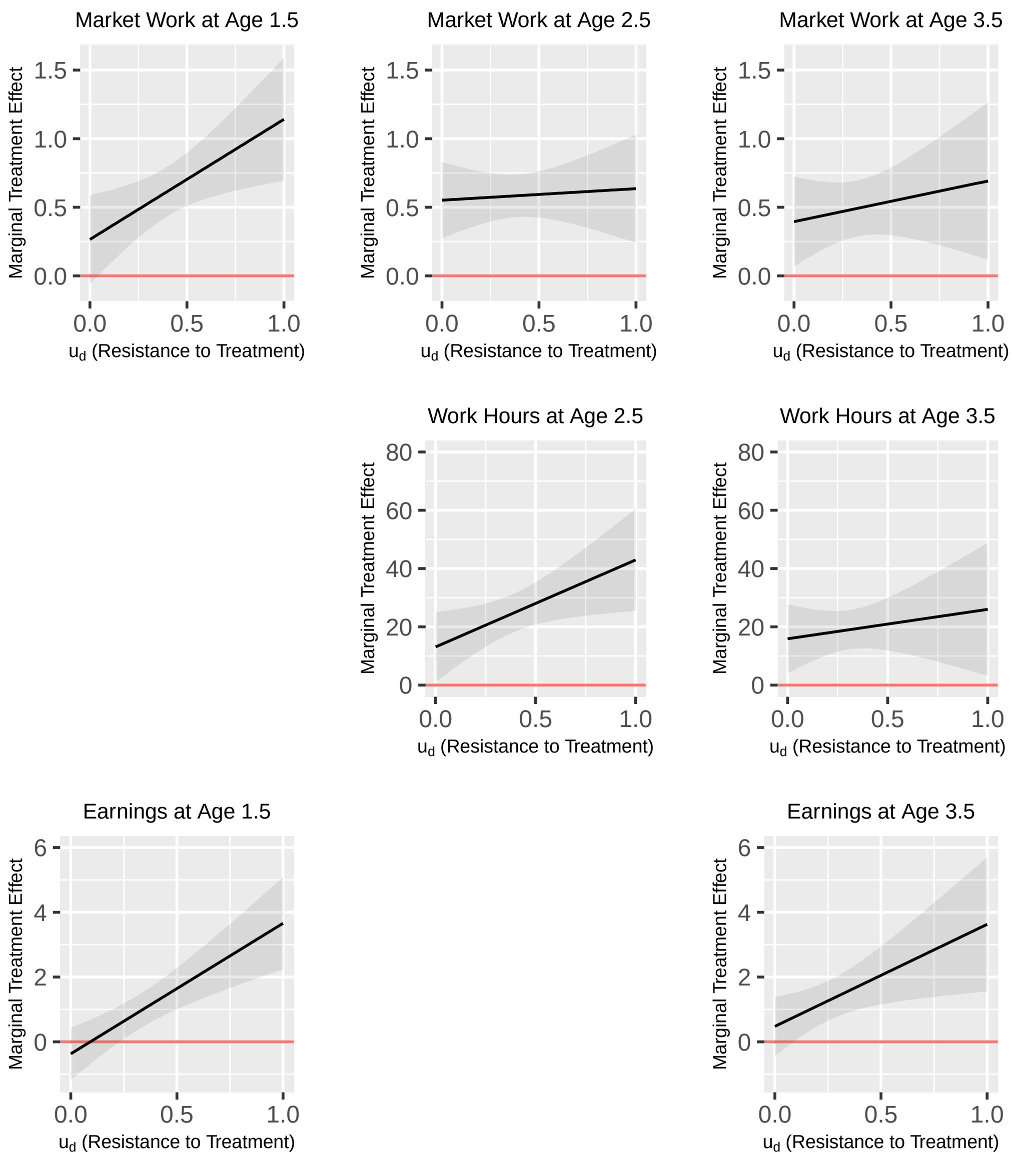

Figure 3: Marginal Treatment Effect Curve

Note: The MTE is averaged over observed characteristics and shown with the 90% confidence interval. Standard errors are clustered at the region level and estimated by a bootstrap with 100 replications. The dependent variables are employment (0 or 1), weekly hours of work, and annual earnings in million JPY. The explanatory variables used in the local IV estimator include the propensity score up to the second-order term, parents' age and education, regional characteristics in 2000 interacted with year dummies, and sets of year and region dummies. To allow for heterogeneous treatment effects by observed characteristics, we interact the propensity score with parents' characteristics.

tract, while a typical nonregular worker is not. The MTE on regular work is positive and significant for a broad range of the resistance to treatment, while the MTE on nonregular work is positive and marginally insignificant. We find no effects on self-employment. These graphs indicate that childcare use increases mothers' market work largely through regular work instead of nonregular work or self-employment.

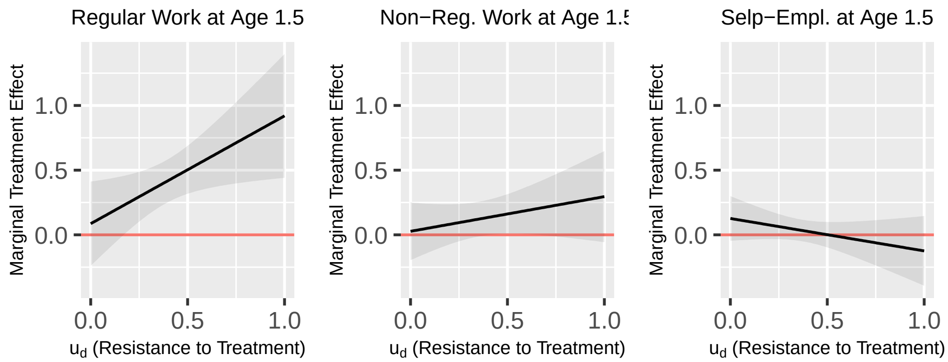

Figure 4: Marginal Treatment Effect Curve on Employment Type
Note: The MTE is averaged over observed characteristics and shown with the 90% confidence interval. See the note for Figure 3 for details.

The MTE also varies by observed characteristics as shown in Table 7. The treatment effect increases with mothers' age for most outcomes. The effects on work and work hours are smaller for mothers with low (less-than-high-school) or high (four-year university or above) education than a mother with medium education (high school or two-year college education). A possible explanation is that many low- and high-education mothers have alternative informal childcare options. In LSN21, 25% of low-education mothers live with their parents or in-laws, while 20% of medium-education mothers do so. This implies that informal childcare by grandparents is more available for low-education mothers than for medium-education mothers. Fewer high-education mothers live with their parents or in-laws, but they may have a greater willingness to pay for informal childcare services or to seek help from their parents. For both low- and high-education mothers, the provision of childcare is more likely to cause crowding out of their informal care arrangement than it is for medium-education mothers.

While the MTE on participation and hours is nonlinear in mothers' education, the MTE on earnings increases largely with mothers' education. This is because the hourly wage increases with mothers' education, which dominates the effects on participation and hours. A similar difference by mothers' education is also seen for employment type. The MTE on regular work increases with education, while that on nonregular work decreases with education.

The MTEs tend to decrease with father's age and education, although some estimates are noisy. Because the primary earner is typically the father in Japan, mothers' labor supply decreases with father's earnings. When father's earnings are low, the mother tries hard to find informal childcare and to work to raise her family. Providing a childcare place to this mother is likely to crowd out her informal childcare without affecting her labor market outcomes.

Aggregate treatment effect parameters including the ATE, TT, and TUT can be calculated by taking a weighted average of the MTE. The weights for the aggregate treatment effect parameters are graphically presented in Figure 5 when the child is aged 1.5 years (see Appendix A for details). For the TT, greater weight is given to those with lower values of resistance to treatment. In contrast, for the TUT, greater weight is given to those with higher values of resistance to treatment. The weights are similar for other age groups.

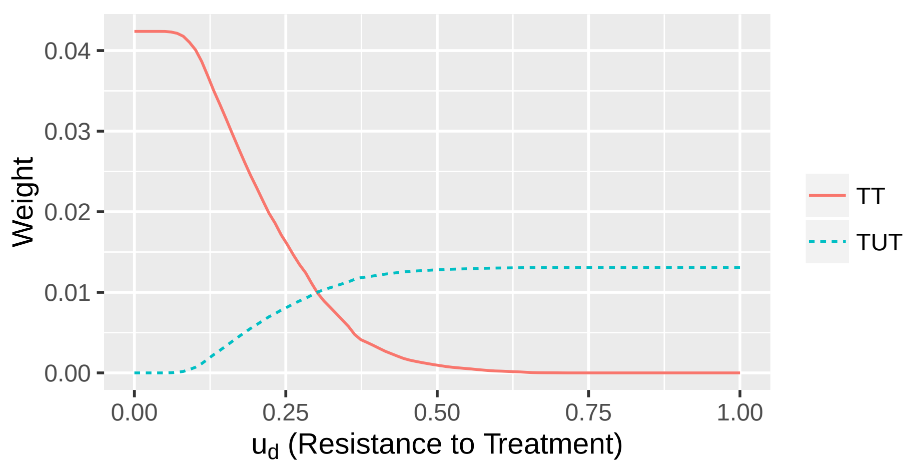

Figure 5: Weights for Treatment Effects on the Treated and Untreated
Note: Weights for TT and TUT for mothers of children age 1.5 are presented. The details for weights are given in Appendix A.

Table 7: Treatment Effect Heterogeneity by Observed Characteristics

	Market Work			**Work Hours**		**Earnings**		**Reg. Work**	**Non-Reg. Work**	**Self-Empl.**
	Age 1.5	Age 2.5	Age 3.5	Age 2.5	Age 3.5	Age 1.5	Age 3.5	Age 1.5	Age 1.5	Age 1.5
Mother										
Age	0.009	0.012	0.008	0.503	0.548	0.045	0.039	0.013	-0.007	0.002
	(0.005)	(0.004)	(0.004)	(0.160)	(0.175)	(0.015)	(0.013)	(0.004)	(0.003)	(0.002)
Less Than HS	0.018	-0.042	-0.081	-3.025	-8.142	-0.190	-0.663	-0.146	0.186	-0.001
	(0.110)	(0.080)	(0.077)	(3.571)	(2.732)	(0.166)	(0.218)	(0.063)	(0.069)	(0.055)
HS	0.161	0.136	0.143	2.693	0.209	-0.191	-0.558	-0.138	0.259	0.041
	(0.066)	(0.059)	(0.056)	(2.234)	(2.048)	(0.168)	(0.202)	(0.060)	(0.061)	(0.024)
2-Yr College	0.149	0.169	0.208	5.715	3.970	0.011	-0.095	0.014	0.135	0.021
	(0.072)	(0.047)	(0.051)	(1.824)	(1.802)	(0.171)	(0.194)	(0.062)	(0.051)	(0.022)
Father										
Age	-0.007	-0.005	-0.005	-0.298	-0.319	0.001	-0.017	-0.005	-0.002	0.001
	(0.004)	(0.003)	(0.002)	(0.155)	(0.117)	(0.013)	(0.012)	(0.004)	(0.003)	(0.002)
Less Than HS	-0.101	-0.044	-0.268	-3.792	-4.133	-0.466	-1.082	-0.039	-0.028	0.028
	(0.080)	(0.086)	(0.114)	(3.192)	(4.022)	(0.177)	(0.342)	(0.063)	(0.066)	(0.040)
HS	-0.057	0.003	-0.116	-0.150	-1.277	-0.354	-0.839	-0.029	0.024	-0.003
	(0.049)	(0.055)	(0.077)	(2.397)	(2.654)	(0.144)	(0.247)	(0.048)	(0.041)	(0.020)
2-Yr College	0.086	0.016	-0.047	2.163	2.099	-0.209	-0.475	0.100	0.032	-0.014
	(0.060)	(0.050)	(0.078)	(1.991)	(2.886)	(0.193)	(0.276)	(0.066)	(0.039)	(0.029)

Note: Standard errors are in parenthesis. They are clustered at the region level and estimated by a bootstrap with 100 replications. The dependent variables are employment (0 or 1), weekly hours of work, and annual earnings in million JPY. The explanatory variables used in the local IV estimator include the propensity score up to the second-order term, parents' age and education, regional characteristics in 2000 interacted with year dummies, and sets of year and region dummies. To allow for heterogeneous treatment effects by observed characteristics, we interact the propensity score with parents' characteristics.

Table 8 presents estimates for the aggregate treatment effect parameters. The ATE on mothers' work ranges from 0.543 to 0.703 and decreases with child's age. For all age groups, the TT on mothers' work is smaller than the TUT, which implies negative selection into treatment. This result is consistent with the upward-sloping MTE curves seen in Figure 3. The analysis here takes into account not only unobserved but also observed heterogeneity and hence gives a more complete picture of the relationship between the treatment effect and selection pattern.

The ATE on weekly hours of work is positive and significant and decreases with child's age, while the ATE on annual earnings is positive and significant, and increases with child's age. For all of these outcomes, the TT is smaller than the TUT, and the differences are statistically significant except for work hours when the child is at age 3.5 years.

ATE on regular work is about three times greater than ATE on nonregular work, while they are both positive and significant. The TT is smaller than the TUT for both types of work. We find no effect on self-employment.

7 Discussion

7.1 Robustness Checks

We examine the robustness of our main results from different perspectives. The first issue is endogenous fertility. The identifying variation of the causal effects is the regional variation of the childcare coverage rate, which is given by the number of childcare places per child. Because the coverage rate is a measure of childcare availability, it may also affect fertility and hence the coverage rate itself. To avoid this concern from endogenous fertility, we use an alternative instrument given by the number of childcare place per woman aged 20–44 years. Arguably, this instrument is more likely to be exogenous than the coverage rate.

The second issue is selective migration. If mothers with stronger labor market attachment move to a region where childcare is more readily available, our estimates are upward biased. To assess the extent of biases from selective migration, we estimate the model using the coverage rate and other regional characteristics at the

Table 8: Aggregate Treatment Effects

	Market Work			**Work Hours**		**Earnings**		**Reg. Work**	**Non-Reg. Work**	**Self-Empl.**
	Age 1.5	Age 2.5	Age 3.5	Age 2.5	Age 3.5	Age 1.5	Age 3.5	Age 1.5	Age 1.5	Age 1.5
ATE	0.703	0.594	0.543	28.029	20.946	1.644	2.052	0.503	0.161	0.001
	(0.118)	(0.103)	(0.150)	(4.425)	(5.487)	(0.390)	(0.544)	(0.113)	(0.093)	(0.060)
TT	0.386	0.567	0.455	18.467	17.950	0.176	1.113	0.201	0.064	0.092
	(0.153)	(0.121)	(0.138)	(5.110)	(4.553)	(0.384)	(0.378)	(0.147)	(0.106)	(0.077)
TUT	0.801	0.606	0.595	32.455	22.698	2.087	2.595	0.596	0.191	-0.027
	(0.142)	(0.133)	(0.209)	(5.855)	(8.091)	(0.476)	(0.771)	(0.143)	(0.114)	(0.079)
TT - TUT	-0.415	-0.039	-0.139	-13.988	-4.748	-1.911	-1.482	-0.395	-0.127	0.119
	(0.194)	(0.168)	(0.227)	(7.430)	(9.010)	(0.556)	(0.763)	(0.210)	(0.145)	(0.118)

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

province level, which is a more aggregated level than the preferred specification. Evidence from the Employment Structure Survey 2012 shows that 9.5% of mothers of children under 6 years of age moved within the same city for childrearing and education in their last move, but only 1.4% moved from other provinces. Hence, we have much less concern about biases from selective migration in this alternative specification.

The third issue is the role of siblings. We do not include the number of siblings in the preferred specification, because fertility may be affected by the availability of childcare. However, if it is exogenous and relevant for childcare enrollment and/or mothers' labor supply, controlling for the number of siblings can reduce the size of the standard errors. Indeed, municipal governments prefer an application from a family in which an older sibling has already attended childcare. We estimate the propensity score and the outcome equations using a model that includes the numbers of younger and older siblings.

The fourth issue is concerned with the model for childcare enrollment. In our preferred specification, we use the coverage rate in the current year only, but childcare enrollment may also be affected by the coverage rates in previous years. This is because some children started childcare at an earlier age, and they can remain enrolled if they wish. The enrollment status of these children is not affected by the current coverage rate but by the coverage rates in previous years. Although leaving out the coverage rates in previous years does not bias our estimates, the estimates may become more precise by using these additional variables. We estimate the propensity score using a probit model augmented by the second-order polynomials of the coverage rates from age 0.5 years to the current age.

The fifth and last issue is model specification. While we assume that the MTE changes linearly with the unobserved resistance to treatment u_D, this assumption may be restrictive. We allow for nonlinearity of the MTE curve by including polynomial terms up to the fourth order in the outcome equations.

We assess the sensitivity of our main results by comparing the estimates of the aggregate treatment effect parameters in the benchmark model with those of the alternative models. Table 9 reports the estimates for selected outcomes. Our estimates are largely robust to the issues raised above, although statistical significance may change across specifications. The estimates are significantly different from the baseline specification when the fourth-order term of the propensity score is included in the local IV regression. The point estimates do not seem to be reasonable, and their standard errors are large, which suggests that the fourth-order term is irrelevant for outcomes, and including it only increases noise in the estimates. Overall, our main results seem to be robust to the endogenous fertility, selective migration, and model specification issues.

7.2 Interpretation

Our analysis indicates that mothers with weaker treatment effects are more likely to be treated, while mothers with stronger treatment effects are less likely to be treated. To understand why selection has a negative effect on the treatment effect, consider how childcare enrollment is determined. Families decide on whether they apply for a childcare center given the pecuniary and nonpecuniary benefits of the childcare use relative to nonuse. If there are more applications than available places, the local governments rank their applications by how many hours their parents work when they use a childcare center. This ranking corresponds to the unobserved resistance to treatment u_D in our model. Hence, u_D should be interpreted as (the negative of) the unobserved component of the preference for work.

Why is the MTE negatively associated with the unobserved preference for work? The treatment effect is the difference in an outcome (e.g., labor market participation) between the treated and untreated states. In the treated state in which families use childcare, most mothers work because it is effectively a prerequisite for childcare use. However, in the nontreated state, the probability of labor market participation may vary considerably by individual depending on the availability and affordability of alternative childcare arrangements.

Mothers with a strong preference for work (or low u_D) are likely to make an extra effort to find an informal childcare arrangement. For example, young families may choose to live close to their parents or in-laws so that grandparents can take care of the children (see Compton and Pollak (2014)). Seeking help from in-laws may not be painless depending on the relationship, given that old Japanese people

Table 9: Robustness of Main Results

	(1) Baseline	(2) Endogenous Fertility	(3) Migration	(4) Siblings	(5) Lagged Cov. Rate	(6) 3rd-Order	(7) 4th-Order
Market Work at Age 1.5							
ATE	0.703	0.675	0.703	0.607	0.691	0.793	1.573
	(0.118)	(0.123)	(0.180)	(0.097)	(0.135)	(0.481)	(1.320)
TT	0.386	0.473	0.692	0.258	0.374	0.433	0.255
	(0.153)	(0.170)	(0.293)	(0.143)	(0.148)	(0.254)	(0.270)
TUT	0.801	0.737	0.706	0.715	0.789	0.904	1.981
	(0.142)	(0.159)	(0.218)	(0.118)	(0.170)	(0.567)	(1.762)
TT-TUT	-0.415	-0.264	-0.014	-0.457	-0.415	-0.471	-1.726
	(0.194)	(0.246)	(0.367)	(0.181)	(0.219)	(0.403)	(1.875)
Work Hours at Age 2.5							
ATE	28.029	22.822	38.423	41.979	27.520	32.728	27.877
	(4.425)	(4.078)	(5.552)	(3.482)	(4.683)	(6.759)	(11.419)
TT	18.467	13.739	35.371	22.371	16.700	22.760	27.280
	(5.110)	(5.391)	(5.627)	(4.636)	(5.494)	(7.325)	(10.786)
TUT	32.455	27.027	39.836	51.060	32.529	37.342	28.154
	(5.855)	(5.587)	(6.883)	(4.130)	(5.582)	(7.653)	(18.869)
TT-TUT	-13.988	-13.288	-4.465	-28.689	-15.829	-14.582	-0.874
	(7.430)	(8.039)	(7.257)	(5.693)	(6.389)	(7.330)	(26.222)
Earnings at Age 1.5							
ATE	1.644	1.673	1.672	1.474	1.630	1.040	2.933
	(0.390)	(0.446)	(0.371)	(0.294)	(0.417)	(0.922)	(3.266)
TT	0.176	0.551	0.603	-0.072	0.141	-0.138	-0.568
	(0.384)	(0.465)	(0.473)	(0.358)	(0.404)	(0.543)	(0.617)
TUT	2.087	2.011	1.994	1.940	2.079	1.394	3.988
	(0.476)	(0.517)	(0.429)	(0.345)	(0.510)	(1.082)	(4.353)
TT-TUT	-1.911	-1.460	-1.390	-2.012	-1.938	-1.532	-4.556
	(0.556)	(0.558)	(0.557)	(0.449)	(0.588)	(0.802)	(4.702)

Note: Standard errors are in parenthesis and clustered at the region level. The ATE is the average treatment effect, TT is the treatment effect on the treated, and TUT is the treatment effect on the untreated. Model (1) is the baseline model. In model (2), the instrument is the number of childcare slots per women aged 20-44. In model (3), regional variables are aggregated to the province level. In model (4), the numbers of younger and older siblings are included. In model (5), the coverage rates in previous years since age 0.5 are also included in the set of instruments. In models (6) and (7), up to the third and fourth order polynomials of the propensity score are included, respectively.

tend to have the traditional family value that the mother should stay at home to raise her children. However, a mother with a stronger preference for work may not hesitate to ask in-laws to take care of her children, because she is desperate for childcare. This mother is likely to work even in the untreated state in which she is not given a childcare place. This implies that the treatment effect is small for those with low u_D.

In contrast, mothers with a weak preference for work are unlikely to exert much effort to find an informal childcare arrangement in the untreated state. They are willing to work only if they are given a place in a childcare center. This implies that the treatment effect is large for those with high u_D.

7.3 Policy Simulations

Using the estimated model, we simulate childcare reforms that increase the coverage rate. These reforms do not change the distribution of the treatment effects, but their policy effects vary because different policies induce different individuals into treatment.

In the first simulation, we evaluate a policy that changes the coverage rate from 0.28 to 0.35, which corresponds to the change from 2002 to 2011. This simulation is useful for understanding the effects of childcare expansion during this period. In the second simulation, the coverage rate is raised from the 2011 level (= 0.35) to 0.42. The size of the change in the coverage rate is the same as that in the first simulation. In the third simulation, the coverage rate is further increased from 0.42 to 0.82, which is the highest coverage rate in the sample.

Policies are evaluated by aggregating the MTE to the policy-relevant treatment effect (see Heckman and Vytlacil (2005)). Suppose that a new policy changes the propensity score for an individual i from p_i to p_i'. Let $\bar{p}$ and $\bar{p}'$ be the sample means of the propensity score in the baseline policy and a new policy, respectively. The policy-relevant treatment effect $PRTE$ is given by

$$PRTE = \frac{1}{N}\sum_{i}^{N} \frac{p_i' - p_i}{\bar{p}' - \bar{p}} MTE_i,$$

where N is the number of individuals in the sample, and MTE_i is the MTE for individual i. Details are provided in Appendix A.

The weights for policy-relevant treatment effects are graphically presented in Figure 6. As the coverage rate increases from one policy to another, greater weight is given to individuals with a higher unobserved resistance to treatment u_D. This implies that individuals with a higher resistance to treatment are induced gradually into treatment, as childcare reforms progress.

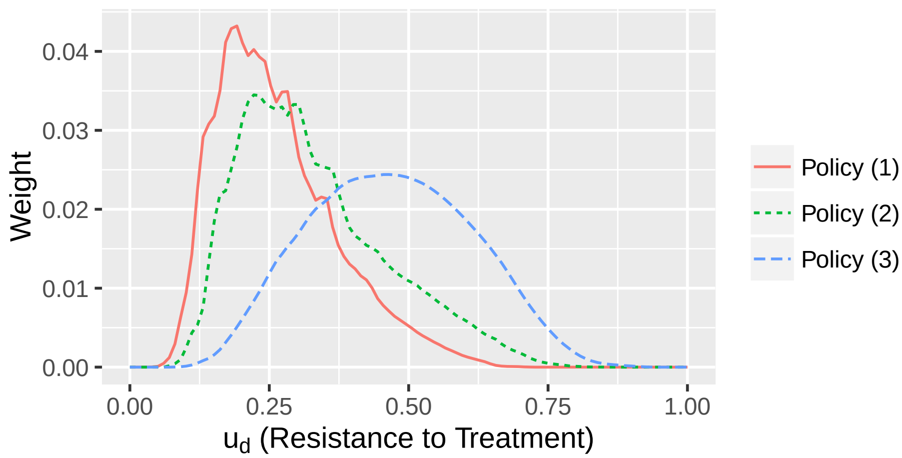

Figure 6: Weights for Policy-Relevant Treatment Effects

Note: In policy (1), the coverage rate is increased from 0.28 to 0.35. In policy (2), the coverage rate is increased from 0.35 to 0.42. In policy (3), the coverage rate is increased from 0.42 to 0.82.

Table 10 summarizes the results. The childcare expansion during 2002–2011 increased the coverage rate from 0.28 to 0.35, which eventually changed the childcare enrollment or propensity score from 0.214 to 0.257 for children aged 1.5 years. The corresponding policy-relevant treatment effects on mothers' work and earnings are 0.497 and 0.693, respectively. If the coverage rate continues to increase by the same extent from 0.35 to 0.42, the enrollment rate increases from 0.257 to 0.306. While the change in the enrollment rate is similar to that in the first simulation, the policy-relevant treatment effects on work and earnings are 0.554 and 0.956, respectively, which are greater than those in the first simulation.

A further childcare expansion has even stronger effects. Raising the coverage rate from 0.42 to 0.82 increases the enrollment rate to 0.632 for children aged 1.5

years, which is close to the enrollment rate for children aged 0–2 years in formal childcare in Denmark, where the participation rate is highest among the OECD countries (see NOSOSCO (2015)). The policy-relevant treatment effects of the third simulation are 0.676 for work and 1.517 for earnings, which are greater than those of the first two simulations.

The simulation results for children aged 2.5 and 3.5 years are also reported in Table 10. For both age groups, the simulated policies increase childcare enrollment, mothers' work, hours of work, and earnings. The policy-relevant treatment effects become increasingly stronger as the coverage rate increases.

Table 10: Counterfactual Simulations of Policies to Increase the Coverage Rate

	Propensity Score		**Policy-Relevant Treatment Effect**		
	Baseline	New Policy	Work	Hours	Earnings
Age 1.5					
(1) Raise Coverage Rate from 0.28 to 0.35	0.214	0.257	0.497		0.693
	(0.010)	(0.014)	(0.114)		(0.340)
(2) Raise Coverage Rate from 0.35 to 0.42	0.257	0.306	0.554		0.956
	(0.014)	(0.026)	(0.110)		(0.339)
(3) Raise Coverage Rate from 0.42 to 0.82	0.306	0.632	0.676		1.517
	(0.026)	(0.108)	(0.133)		(0.444)
Age 2.5					
(1) Raise Coverage Rate from 0.28 to 0.35	0.282	0.340	0.580	23.141	
	(0.012)	(0.014)	(0.084)	(3.874)	
(2) Raise Coverage Rate from 0.35 to 0.42	0.340	0.401	0.586	25.501	
	(0.014)	(0.021)	(0.076)	(3.785)	
(3) Raise Coverage Rate from 0.42 to 0.82	0.401	0.786	0.600	30.326	
	(0.021)	(0.072)	(0.086)	(4.620)	
Age 3.5					
(1) Raise Coverage Rate from 0.28 to 0.35	0.326	0.389	0.499	19.426	1.578
	(0.014)	(0.013)	(0.133)	(4.479)	(0.482)
(2) Raise Coverage Rate from 0.35 to 0.42	0.389	0.437	0.514	19.938	1.738
	(0.013)	(0.020)	(0.139)	(4.912)	(0.543)
(3) Raise Coverage Rate from 0.42 to 0.82	0.437	0.562	0.538	20.771	1.998
	(0.020)	(0.111)	(0.161)	(6.211)	(0.681)

Note: Standard errors are in parenthesis. They are calculated by bootstrap with 100 replications and clustered at the region level.

Overall, our simulations indicate that the policy-relevant treatment effects become increasingly stronger as childcare reforms progress. Mothers with weak labor market attachment would be strongly affected by the treatment, but they are unlikely to use a childcare center, because the rationing rule gives them a lower rank. As the coverage rate increases, mothers with weak labor force attachment gradually start to use a childcare center, which explains why the treatment effects become increasingly stronger. Our analysis suggests that the current rationing rule causes inefficient allocation of mothers' labor supply is concerned. Even if the government is unable to use a price mechanism for political ideologies, not ranking applications by how much parents work may improve the efficacy of a childcare reform.

8 Conclusion

We estimate the MTE of childcare enrollment on mothers' labor market outcomes by exploiting regional variations in the growth of the childcare coverage rate. The demand for subsidized childcare exceeds the supply in many regions, and the local governments rank childcare applications by how much parents work in order to assign a childcare place to families in need. Mothers' labor force attachment is a key determinant for a successful application, but it is not observed in the data. The MTE framework enables us to estimate variation in the treatment effects by unobserved propensity for treatment.

Our estimates indicate that mothers with stronger MTE are less likely to use childcare, whereas mothers with weaker MTE are more likely to find a childcare place. The rationing rule prioritizes mothers with stronger labor market attachment, but they are likely to exert extra effort to find an alternative informal childcare arrangement if they are not given a childcare place. Because these mothers are likely to work regardless of the availability of subsidized childcare, the treatment effects on these mothers are small. Mothers with weak labor market attachment are unlikely to work without subsidized childcare, which implies that the treatment effects on these mothers are strong. As long as mothers' labor market outcomes are concerned, our analysis suggests that the current rationing rule may be inefficient.

We also find significant treatment effect heterogeneity by the child's age. The estimates show that the treatment effect on mothers' labor market outcomes decrease with the child's age, which is robust to alternative modeling assumptions. However, note that this result does not necessarily imply that a childcare place should be assigned to mothers of infants, because infant care is particularly expensive.

The main limitation of this paper is that we have not identified an optimal rationing rule and only suggested a potential problem with the current rationing rule. An optimal rationing rule should take into account not only the mothers' labor market outcomes but also child development[11] and other dimensions of family welfare in the long run. We leave this important research agenda for future work.

References

ANDRESEN, M. E. AND T. HAVNES (2016): "Child Care and Parental Labor Supply: A New Look," University of Oslo.

AOBA WARD, CITY OF YOKOHAMA (2009): "Guidance for Enrollment in Childcare in 2010," Yokohama.

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

ASAI, Y., R. KAMBAYASHI, AND S. YAMAGUCHI (2015): "Childcare Availability, Household Structure, and Maternal Employment," *Journal of the Japanese and International Economies*, 38, 172–192.

——— (2016): "Crowding-Out Effect of Subsidized Childcare," McMaster University.

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.

BAUERNSCHUSTER, S. AND M. SCHLOTTER (2015): "Public Child Care and Mothers' Labor Supply-Evidence from Two Quasi-Experiments," *Journal of Public Economics*, 123, 1–16.

BERLINSKI, S. AND S. GALIANI (2007): "The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment," *Labour Economics*, 14, 665 – 680.

[11] See Yamaguchi, Asai, and Kambayashi (2017).

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.

BRINCH, C., M. MOGSTAD, AND M. WISWALL (forthcoming): "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*.

CABINET OFFICE (2010): "A Project for Eliminating Wait-Listed Children by the National and Municipality Governments," available from http://www.kantei.go.jp/jp/singi/taikijidou.

CASCIO, E. U. (2009): "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools," *Journal of Human Resources*, 44, 140–170.

COMPTON, J. AND R. A. POLLAK (2014): "Family Proximity, Childcare, and Women's Labor Force Attachment," *Journal of Urban Economics*, 79, 72–90.

CORNELISSEN, T., C. DUSTMANN, A. RAUTE, AND U. SCHÖNBERG (2016): "From LATE to MTE: Alternative Methods for The Evaluation of Policy Interventions," *Labour Economics*, 41, 47 – 60, sOLE/EALE Conference Issue 2015.

FITZPATRICK, M. D. (2010): "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten," *Journal of Labor Economics*, 28, 51–85.

——— (2012): "Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool," *Journal of Human Resources*, 47, 583–612.

GELBACH, J. B. (2002): "Public Schooling for Young Children and Maternal Labor Supply," *American Economic Review*, 92, 307–322.

GOUX, D. AND E. MAURIN (2010): "Public school availability for two-year olds and mothers' labour supply," *Labour Economics*, 17, 951–962.

HAVNES, T. AND M. MOGSTAD (2011): "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 95, 1455–1465.

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.

KAMBAYASHI, R. AND T. KATO (2013): "Good Jobs, Bad Jobs, and the Great Recession: Lessons from Japan's Lost Decade," Center on Japanese Economy and Business Working Papers 326, Columbia University.

LEFEBVRE, P. AND P. MERRIGAN (2008): "Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada," *Journal of Labor Economics*, 26, 519–548.

LUNDIN, D., E. MÖRK, AND B. ÖCKERT (2008): "How Far Can Reduced Childcare Prices Push Female Labour Supply?" *Labour Economics*, 15, 647–659.

MINISTRY OF HEALTH, LABOUR AND WEALTH (2008): "Basic Policy Stance for Making a New Policy Scheme to Support the Development of the Next Generation," Available from http://www.mhlw.go.jp/shingi/2008/05/s0520-6.html.

——— (2009): "Summary of the Survey of Regional Child Welfare Services 2009," Available from http://www.mhlw.go.jp/toukei/saikin/hw/jidou/09/dl/80.pdf.

NISHITATENO, S. AND M. SHIKATA (2017): "Has Improved Daycare Accessibility Increased Japan's Maternal Employment Rate? Municipal Evidence from 2000-2010," *Journal of the Japanese and International Economies*, 44, 67–77.

NOLLENBERGER, N. AND N. RODRIGUEZ-PLANAS (2015): "Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain," *Labour Economics*, 36, 124 – 136.

NOSOSCO (2015): *Social Protection in Nordic Countries*.

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): "How Does Early Childcare Enrollment Affect Children, Parents, and Their Interactions?" McMaster University.

A Treatment Parameters

We calculate treatment parameters following the method outlined by Cornelissen et al. (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 PRTE is given by

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