TCER Working Paper Series

Wholesome Lunch to the Whole Classroom: Short- and Longer-Term Effects on Early Teenagers' Weight

Shiko Maruyama

Sayaka Nakamura

April 2021

Working Paper E-160

https://www.tcer.or.jp/wp/pdf/e160.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH

1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

Previous studies on the effect of school lunch programs on child obesity have been hampered by effect heterogeneity, self-selection, and stigma-induced under-reporting, having produced mixed findings. Their potential long-lasting effect has also been debated. We study the body-weight effect of a Japanese school lunch program, which provides nutritional lunch to all students at participating municipal junior highs. The lack of means testing and individual participation choice offers easily interpretable causal estimates. By exploiting almost all school lunch coverage for elementary school children nationwide, we construct a difference-in-differences (DID) framework to alleviate potential bias due to unobserved differences across municipalities. Using the 1975–1994 National Nutrition Survey, a nationally representative household survey with measured height and weight, we find a regressive benefit of school lunch: while no statistically significant effect is found for the full sample, we find a significant obesity-reducing effect for the subsamples of children with low socioeconomic backgrounds. The obesity-reducing effect remains at least a few years after graduation, implying effect through not only nutritional contents but also guiding healthy eating behavior. We find little evidence that school lunch reduces underweight. Propensity score weighting, quantile DID analysis, and various falsification tests confirm the robustness of our estimates.

Shiko Maruyama
University of Technology Sydney
Economics Discipline Group
PO Box 123, Broadway, NSW 2007 Australia
shiko.maruyama@uts.edu.au

Sayaka Nakamura
TCER
and
Nagoya University
School of Economics
Furocho, Chikusa, Nagoya, 464–8601 Japan
nakamuras@soec.nagoya-u.ac.jp

Wholesome Lunch to the Whole Classroom: Short- and Longer-Term Effects on Early Teenagers' Weight

Shiko Maruyama [a,*], Sayaka Nakamura [b]

[*] Corresponding author

[a] Economics Discipline Group, University of Technology Sydney, PO Box 123, Broadway, NSW 2007 Australia, Email: shiko.maruyama@uts.edu.au

[b] School of Economics, Nagoya University, Furocho, Chikusa, Nagoya, 464–8601 Japan, Email: nakamuras@soec.nagoya-u.ac.jp

THIS VERSION: April 20, 2021

JEL codes: H51, I18, I28

Keywords: school lunch; school nutrition programs; difference-in-differences; child obesity; heterogeneous treatment effect

Wholesome Lunch to the Whole Classroom: Short- and Longer-Term Effects on Early Teenagers' Weight

Abstract (200 words)

Previous studies on the effect of school lunch programs on child obesity have been hampered by effect heterogeneity, self-selection, and stigma-induced under-reporting, having produced mixed findings. Their potential long-lasting effect has also been debated. We study the body-weight effect of a Japanese school lunch program, which provides nutritional lunch to *all* students at participating municipal junior highs. The lack of means testing and individual participation choice offers easily interpretable causal estimates. By exploiting almost all school lunch coverage for elementary school children nationwide, we construct a difference-in-differences (DID) framework to alleviate potential bias due to unobserved differences across municipalities. Using the 1975–1994 National Nutrition Survey, a nationally representative household survey with measured height and weight, we find a regressive benefit of school lunch: while no statistically significant effect is found for the full sample, we find a significant obesity-reducing effect for the subsamples of children with low socioeconomic backgrounds. The obesity-reducing effect remains at least a few years after graduation, implying effect through not only nutritional contents but also guiding healthy eating behavior. We find little evidence that school lunch reduces underweight. Propensity score weighting, quantile DID analysis, and various falsification tests confirm the robustness of our estimates. (199 words)

Highlights

- The Japanese school lunch program is non-means-tested and satisfies strict nutritional requirements.
- We study its effect on junior-high students' weight using a DID framework.
- No statistically significant effect is found for the full sample.

- We find significant obesity-reducing effects for subsamples of low-SES children.
- The obesity-reducing effect remains at least a few years after graduation.

1. Introduction

Child obesity is a serious public health concern that is rapidly growing around the world (Reilly and Kelly, 2011; Ng et al., 2014), and school lunch reforms have recently been implemented to combat this threat. For example, stricter nutritional standards were introduced in the UK and the US during the 2000s and 2010s, respectively, concurrent with the continuing expansion of eligibility for free school lunches in the UK (Evans and Harper, 2009; Baidal and Taveras, 2014). These reforms, however, are under heated debate due to their high costs and the lack of agreement on the effectiveness (Tickle, 2014; Baidal and Taveras, 2014). von Hinke Kessler Scholder (2013) finds the UK school lunch program has had no significant effect. For the US National School Lunch Program (NSLP), some studies find weight-increasing effects (Schanzenbach, 2009; Li and Hooker, 2010; Millimet et al., 2010; Hernandez et al., 2011; Capogrossi and You, 2017), but others find no significant effect (Hofferth and Cuntin, 2005; Gleason and Dodd, 2009; Mirtcheva and Powell, 2013) or find weight-reduction effects (Gundersen et al., 2012).[1]

Differences across programs have most likely contributed to the lack of consensus. Yet, mixed findings, even for a single program, might suggest individual-level effect heterogeneity. In the context of school lunch programs, effect heterogeneity is plausible because the causal effect depends on not only what students eat at school lunch but also what they would have eaten in the absence of the program. School lunch will increase the weight of students who eat little without school lunch and decrease the weight of students who regularly have energy-dense diets such as French fries and sweetened beverages. Smith (2017) finds substantial heterogeneity in the effect of the National

[1] For the US School Breakfast Program (SBP), some studies find weight-reduction effects (Gleason and Dodd, 2009; Millimet et al., 2010; Millimet and Tchernis, 2013), while others find no significant effects (Hofferth and Cuntin, 2005; Sudharsanan et al., 2016).

School Lunch Program (NSLP) on the quality of children's dietary intake, with positive impacts for below-median quantiles of nutritional quality distribution and negative effects for upper quantiles.

Another important debate is the long-lasting impact on obesity that school meal programs may have by altering children's food preferences and eating habits. The nutrition education and encouragement of good eating habits are the stated goal of school lunch programs in many countries, including Italy, Sweden, and Finland (Andersen et al., 2017), Japan (NIEPR, 2013), and South Korea (Jo, 2014). A long-lasting effect, if it exists, that could drastically enhance the cost-effectiveness of a school meal program, and hence it is of policy interest. However, studies on the long-term effect of school meals on weight are scarce and their findings are mixed: Peterson (2014) finds a significant, positive association between NSLP participation and adulthood obesity, though Hinrichs (2010) finds no significant effect of NSLP on adulthood weight.

In this paper, we study the body-weight effect of a Japanese school lunch program for students at municipal junior highs (public lower secondary schools for 12- to 15-year-olds). Unlike the school meal programs previously studied, the Japanese program requires *all students* at participating schools to eat school lunch. We use individual-level data drawn from the 1975–1994 National Nutrition Survey (NNS) and estimate the effect of the school lunch program on measures of weight-for-height, obesity, and underweight, including the body mass index (BMI, defined as [weight in kilograms] / [height in meters]^2). This paper advances the literature in several ways. First, by conducting subsample and quantile analysis, we examine heterogeneity in the effect of school lunch on weight by socioeconomic status (SES) for the first time in the literature. Knowledge of individual-level effect heterogeneity aids policymakers in the design of effective and efficient school meal programs. Second, we estimate not only the contemporaneous effect but also the lagged effect of school meal programs for children who graduate from junior highs and then stop having school lunch.

The unique features of the Japanese program allow us to obtain credible and easily interpretable causal estimates. First, all children attending a participating school must eat the provided lunch. The lack of individual choice in school lunch participation, together with municipality-level variation in school lunch provision, enables us to estimate the population effect of actual participation in the school lunch program and assess treatment heterogeneity by children's socioeconomic background. This is in contrast to the aforementioned existing studies that examine school lunch programs in which participation is up to each family. To address the concern about selection bias, existing studies use regression discontinuity design (RDD) based on the eligibility cutoff for school lunch subsidies (Schanzenbach, 2009) and difference-in-differences (DID) analysis based on policy changes resulting in the increased out-of-pocket cost of school lunch for households in specific income brackets (von Hinke Kessler Scholder, 2013). Consequently, these estimates are a local average treatment effect (LATE) for specific income/SES groups at the margin affected by the policy change or policy discontinuity. Moreover, these estimates are interpreted as the effect of "intention to treat" (ITT) rather than the actual participation effect because of voluntary participation. While a LATE parameter is informative for a marginal policy change, the estimates of population average effect are informative for the overall effect of the program.

The second advantage of studying the Japanese school lunch program is its strict national nutritional standards, which minimize program heterogeneity within the country. The strict standards also make our estimates topical and informative to policymakers in other countries where nutritional requirements are currently debated or have recently been strengthened. Third, there are no other large-scale food provision programs in Japan. In the US there are various means-tested food assistance programs, including NSLP, School Breakfast Program (SBP), and the Food Stamp Program (Hofferth and Cuntin, 2005), and this coexistence of multiple programs complicates the analysis (Capogrossi and You, 2017). Fourth, the non-means-tested nature of the program implies no stigma-induced reporting bias in school lunch participation. This is in contrast to programs in many other countries. Gundersen et al. (2012) show that participation in NSLP is significantly underreported, which is overlooked in other studies.

At the core of our identification strategy is a large variation in the availability of school lunch across municipalities. This municipality-level variation is useful because the program participation decision is up to each municipality rather than the policy agenda of the central government. However, our causal estimate may be confounded if municipalities' decisions reflect unobserved area characteristics. To address this concern, we employ a novel DID framework, in which we exploit elementary school children as pre-treatment data. Japanese compulsory education consists of six years of elementary school for 6- to 12-year-olds and three years of junior high school for 12- to 15-year-olds, where the majority attend municipal schools (Ministry of Education, 1994). In contrast to a large variation in municipal school lunch provision for junior highs, nearly all municipalities provide school lunch for elementary schools (NIEPR, 2013). Exploiting this structure, we form a DID framework in which we compare differences between junior-high students and older (9- to 12-year-old) elementary students between municipalities with and without school lunch for junior highs. Because similarity between the treatment and control municipalities provides further assurance for the common trend assumption of the DID approach, we also undertake DID estimation combined with propensity-score trimming and inverse probability of treatment weighting (IPTW). We further conduct the raw quantile DID (QDID) analysis, i.e., the raw DID using percentiles instead of the means, to examine the effect of school lunch on the distribution of weight measures.

Our main findings are as follows. While we find no significant effect of school lunch on body weight for the entire sample, we find a significant obesity-reducing effect for the subsample of children with low socioeconomic backgrounds. We find little evidence that school lunch reduces underweight prevalence. Quantile DID analysis reveals that the weight-reduction effect is concentrated among obese children. The weight-reduction effect for low-SES children persists several years after they graduate from junior high, implying effect through not only nutritional content but also guiding healthy eating behavior. Various sensitivity tests confirm the robustness of our DID estimates.

2. Backgrounds

School meal programs vary considerably across countries. Historically, dating back to the charitable provision of free lunch for poor and undernourished schoolchildren in the mid-19th century in the UK and US, the original goal of school meal programs was to feed children in hunger. By the mid-20th century, national-level school lunch programs existed in Finland, France, Italy, Ireland, Portugal, Spain, and Sweden, with the major nutritional goal of increasing energy intake rather than reducing it (Evans and Harper, 2009; Rutledge, 2015). Nowadays there is a large variation across developed countries in nutritional content, targeting, and policy goals of school meal programs (Harper et al., 2008).[2]

The scene of school lunch in Japan differs from that in many Western countries. Students at municipal elementary and junior-high schools are divided into a class of about 40 students and spend the entire school day with their classmates, mostly at an assigned seat in the homeroom; students eat lunch with their classmates in the classroom.[3] If a municipality provides lunch, all students have the same menu (Sanborn, 2017).[4] A typical menu consists of a dish of meat or fish, a dish of vegetable, rice or bread, fruit, and milk (Nozue, 2011). Figure 1 shows typical lunch scenes at junior highs. As in Panel (a), students in charge serve the meal. They are instructed to distribute food evenly to all students, but when there are leftovers students can have another serving (Nozue, 2011). Students eat

[2] In developing countries, school feeding programs still play an important role in increasing energy intake and reducing underweight (Jomaa et al., 2011).

[3] A similar setup can be found in Italy, where many elementary and lower-secondary schools ban home-prepared lunches at school so that students will eat school lunch together with their classmates (Kingston, 2019).

[4] Few municipalities offer a menu choice.

together as in Panel (b). The homeroom teacher eats the same school lunch with the students and is supposed to encourage students to eat everything served to them (NIEPR, 2013).

The first government-subsidized school lunch program in Japan started in 1932 for the children at elementary schools from low-income families, but the program was interrupted in 1944 due to World War II. It was then resumed in municipal elementary schools in Tokyo in 1946, when Japan was under American occupation, to combat child malnutrition resulting from severe food shortage. Thereafter the program was gradually expanded to become nationwide. The nutritional and sanitary guidelines became the School Lunch Law (henceforth SLL) in 1954, which was then revised to include municipal junior highs in 1956 (NIEPR, 2013). The SLL obliges municipalities only to “make efforts” to provide lunch at municipal elementary and junior-high schools.

The school lunch coverage rate for municipal elementary school students has been above 98 percent since the late 1970s, but the coverage rate for municipal junior-high students has increased more gradually from about 58 percent in 1978 to about 83 percent in 2015, as detailed in Appendix 1 (Figure A1). Figure 2 shows the 1985 school lunch coverage rate among municipal junior-high students across prefectures, which is evenly distributed across Japan with moderate spatial correlation.

The Ministry of Education[5] requires that all students eat the provided school lunch, and not bring their own food to school, if their school receives government funding for school lunch. The only exception to this requirement is for students with special dietary needs such as a food allergy (Ono, 2007). Nevertheless, households’ ability to pay for school lunch is unlikely to hinder children’s participation. This is because municipalities receive subsidies that cover the cost of construction and maintenance of facilities and the labor costs involved in preparing the food, and the

[5] The Ministry of Education was consolidated into the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) in 2001.

children's guardians only pay ingredient costs and energy bills (NIEPR, 2013). In 2013, for example, the average junior-high lunch fee was about 4,800 yen per person per month (approximately US$43) (MEXT, 2015). Moreover, payments are exempted for households on public assistance, and almost all municipalities provide school lunch to all students regardless of fee payment (Fujisawa, 2008). These compulsory and non-means-tested aspects of the Japanese school lunch program provide a unique opportunity to evaluate the population average treatment effect (ATE) rather than a LATE or ITT effect.[6]

The Ministry of Education has set nutritional standards for school lunch, including target values for energy, protein, total fat, calcium, and vitamin since 1954 (Nozue, 2011), and during our study period the standards have had only one minor revision (see Appendix 1). Japan's standards are significantly higher than their contemporary counterparts in the US and UK. The current US federal requirements lack target values for protein, total fat, calcium, and vitamin, and there was only a minimum requirement for energy intake until 2012 (Baidal and Taveras, 2014). In the UK, requirements for the amount of energy, protein, and fat for school lunch were introduced in 1966 but were abolished in 1980, leaving no legally binding nutrition requirements until 2000 (Evans and Harper, 2009). A survey conducted by the Japanese Ministry of Education finds that the actual intake from school lunch complies well with the nutritional standards (National School Health Center of Japan, 1990, 1991).

The SLL also lists the *educational goals* of the school lunch program such as encouraging good eating habits, imparting knowledge on nutrition, and fostering sociability. The school lunch program has been a part of the curriculum since 1958, instructing teachers to guide table manners and encourage healthy eating (NIEPR, 2013). Since 2005, the school lunch program has been

---

[6] An increasing number of municipalities nowadays allow students to choose between school lunch and home-prepared lunch, but such a choice was rare during our study period (see Appendix 1).

embedded as a part of the national food and nutrition education program called *shokuiku*, which aims to promote healthy eating and prevent obesity and underweight.[7]

Despite the low average BMI of the Japanese, the prevalence of obesity-related diseases in Japan is close to that of other developed countries with higher average BMIs (Guariguata et al., 2014). Because the rise in obesity-related health risks starts at a lower BMI level among Asians than among Caucasians, the Japan Society for the Study of Obesity has advocated defining obesity as BMI of 25 or over since 2000, as opposed to the WHO's benchmark of BMI 30 (Kanazawa et al., 2002). BMI and obesity among Japanese children have significantly increased since the late 1970s (Yoshinaga et al., 2010; Maruyama and Nakamura, 2015), raising concerns about child obesity, which is a major risk factor for adult obesity (Togashi et al., 2002).

Underweight is also a serious public health issue among Japanese women, and is associated with various morbidities (Kodama, 2010). Female BMI has consistently decreased in post-war Japan (Maruyama and Nakamura, 2015), and underweight prevalence has significantly increased among young women since the 1970s (Takimoto, 2004). Currently, more than 20 percent of non-pregnant women in their 20s are underweight (a BMI of 18.5 or lower) (MHLW (Ministry of Health, Labour and Welfare), 2016).

We know of only one study examining the effect of the Japanese school lunch program on child weight. Applying a fixed-effects OLS model with a lagged dependent variable to 2006–2015 prefecture-level data, Miyawaki et al. (2019) find a significantly negative effect of junior high school lunch coverage rate on overweight prevalence for boys but a nonsignificant effect for girls, and a nonsignificant effect on underweight prevalence for both genders. However, not only the lack of time-variant control variables in their model could induce bias but also this type of dynamic model

[7] The Third Basic Program for Shokuiku Promotion: www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000129496.pdf (In Japanese).

makes OLS estimates inconsistent (Nickell, 1981; Angrist and Pischke, 2009, pp. 243–246). Other Japanese studies find a positive association between school lunch and the intake of vegetables and dairy products (Takahashi et al., 1983; Kawaraya and Mori, 2009; Nozue et al., 2010) and the poor nutritional contents of weekday lunch among junior-high students who attend schools without school lunch (National School Health Center of Japan, 1990).

3. Framework for Causal Analysis

3.1. Difference-in-differences (DID) analysis

We conduct DID analysis using data of elementary and junior-high students drawn from the 1975–1994 NNS. The repeated cross-sectional data in the NNS allow for neither tracking the same municipalities/students over time nor identifying the name of the municipality of students' residence. However, encrypted identifiers enable us to group residents by census district, a subarea of a municipality (henceforth "district"), and to impute the school lunch provision status of the municipality the district belongs to.

Without loss of generality, we regard having school lunch as the baseline case and assess the effect of *not* having school lunch, for the ease of presentation. We start with the following linear regression for a sample of junior-high students:

$$Y_{id} = X_{id}\beta + Z_d\gamma + \theta^{OLS} NoSchoolLunch_d + \epsilon_{id} , \tag{1}$$

where Y_{id} denotes an outcome variable (e.g., BMI) of student *i* in district *d*, X_{id} is a vector of student *i*'s characteristics, including a constant term, and Z_d is a vector of district *d*'s characteristics, including dummies for survey years. Because we do not track the same district or student over time, we suppress year subscripts in Eq. (1) for conciseness. $NoSchoolLunch_d$ is an indicator for no junior-high school lunch in district d. ϵ_{id} is the error term, and β and γ are parameters to be estimated. The OLS estimator of the "no school lunch" effect, θ^{OLS}, could be biased due to systematic differences in unobserved local obesogenic environments between districts with and districts without school lunch, such as access to healthy and unhealthy food, urban sprawl, access to parks and sports facilities, and transportation systems (Lake and Townshend, 2006).

To account for potential confounders, we capitalize on a unique feature of the Japanese school lunch program: nearly all municipal elementary schools provide school lunch. This allows us to form a DID framework in which we subtract the pre-treatment difference, i.e., the difference in outcome measures of elementary students between districts with and districts without junior-high school lunch, from the post-treatment difference, i.e., the same difference of junior-high students, in order to adjust for the pre-existing age-invariant difference between the treatment and control districts, as illustrated in Figure 3. Under the age-grade system in Japan, a child's school grade is determined strictly by the child's age on April 2nd. For elementary students, we limit our sample to 9- to 12-year-olds (Grades 4–6) for a better comparison with junior-high students (Grades 7–9). We estimate the following DID equation using the sample of elementary and junior-high students:

$$Y_{id} = X_{id}\beta + \gamma NoSchoolLunch_d + \theta^{DD} JuniorHigh_{id} \times NoSchoolLunch_d + \mu_d + \epsilon_{id}, \quad (2)$$

where $JuniorHigh_{id}$ is an indicator for junior-high students, μ_d denotes district fixed effects, and θ^{DD} is the DID estimator of the effect of not having school lunch. The *JuniorHigh* indicator appears only in the interaction term and not in a separate term, because junior-high attendance is absorbed by age dummies in X_{id}. The model is identified under the standard parallel trend assumption for ϵ_{id} over age between the treatment and control districts, conditional on X_{id}.[8] We use standard errors clustered at the district level to allow for within-district correlation of the error term. Unlike prior quasi-experimental studies, we quantify the effect of *actual* participation rather than the "intention to

[8] We require parallel trends over age within a survey year, whereas standard DID models based on individual- or municipality-level panel data require parallel trends over age for a certain birth cohort or parallel trends over year for a certain age group.

treat" effect, because attending schools with school lunch and participating in the school lunch program are equivalent in the Japanese setup.

3.2. Subsample analysis

We examine effect heterogeneity for low SES children, a group of particular policy interests for two reasons. First, in many countries, including the UK, US, and Canada, the main target of school meal programs is children in low-income families (Harper et al., 2008). Second, we expect larger effects for lower SES children. Low parental SES is associated with children's poor diet quality (Darmon and Drewnowski, 2008), less healthy eating habits (Hazano et al., 2017), and obesity in childhood (Shrewsbury and Wardle, 2008), adolescence (Kachi et al., 2015), and adulthood (Lee, 2013). Smith (2017) finds a larger impact of the US school meal programs on children's nutrition for children with poorer diet quality.

We conduct a subsample analysis for two low-SES groups. Because our data contain no information on parental education or income, we first use paternal occupations as proxies for SES. We study a subsample of children whose fathers are in non-white-collar occupations or unemployed ("children with non-white-collar fathers," henceforth). White-collar occupations refer to white-collar employees and self-employed professionals, such as practicing clinicians, and non-white-collar occupations refer to all the other occupational categories, including non-professional self-employed workers, laborers, and workers in agriculture/fisheries/forestry.[9]

The actual SES of these households, however, may change over time, affecting comparability across children of different ages and hence causing bias in the estimated long-term effects. To address this concern, we study another subsample: children in households whose per-member

[9] We do not include children without fathers in this subsample because we cannot distinguish cases where fathers temporarily live apart from the family for work.

expenditure is below the median ("children with low household expenditure," henceforth).[10] Household expenditure is a reliable measure of SES, capturing material well-being for low-income households (Meyer and Sullivan, 2003).

3.3. Inverse probability of treatment weighting (IPTW) and propensity-score trimming

As a robustness check, we use IPTW and propensity-score trimming in the DID analysis. Combining DID analysis with propensity score methods has been found effective for eliminating potential sources of temporally invariant bias (Smith and Todd, 2005), balancing baseline characteristics between the treatment and control groups and hence making the common trend assumption more plausible. IPTW gives each subject a weight equal to the inverse of the probability of treatment the subject received. Thus, for students in district *d*, the weight equals $\frac{1}{\mathrm{p}_d}$ if *d* is a treatment district and $\frac{1}{1-\mathrm{p}_d}$ if *d* is a control district, where p_d denotes the propensity score of a lack of school lunch in *d*. Propensity-score trimming excludes observations based on the propensity scores to ensure sufficient overlap in characteristics between the treatment and control districts. Further details are provided in Appendix 2.

4. Data

4.1. The National Nutritional Survey

[10] The median is defined among children of the same two-year age group in the same survey year. Because per-member household expenditure in the NNS is reported in intervals that vary by year, we cannot determine which children in the median interval are below the median. To address this, we randomly split the median interval so that all children of the same age group in each year are split into two groups of equal size.

We use a sample of 9- to 15-year-old children in elementary and junior-high schools drawn from the 1975–1994 NNS. The NNS is a nationally representative, cross-sectional survey annually administered by the Ministry of Welfare.[11] Katanoda et al. (2005) confirm the representativeness of the NNS. The response rate, though not reported for all years, appears to be high: among about 5,000 households requested to participate in the 2002 survey, 4,160 participated (MHLW, 2003). The survey includes information on the body measurements, diet, and sociodemographic characteristics of each household member. Height and weight are measured without shoes and with adjustment for the weight of clothes by health professionals and thus are accurate and free from self-reporting bias (Gorber et al., 2007). The elementary and junior-high students consist of 9- to 12-year-olds and 12- to 15-year-olds, respectively. We use the birth month to determine whether 12-year-old children are in an elementary or junior-high school. Children at age 12 in the 1975–1985 surveys are not used because the birth month information is unavailable in these years. A small number of children without a mother and those with missing data are also excluded. Further details of the sample construction are provided in Appendix 3.

4.2. Municipal provision of school lunch

The NNS samples districts from all Japanese prefectures and requests all households in the sampled districts to participate in the survey. Only an encrypted identifier of the district of residence and the name of the prefecture is provided in the publicly available NNS dataset, hence we impute the school lunch provision status of each district using NNS information on whether school children eat school lunch.

[11] The NNS was renamed the National Health and Nutrition Survey in 2003. The Ministry of Welfare was consolidated into the MHLW in 2001.

This procedure involves several measurement issues. The first issue is that school lunch is served only on weekdays, whereas the NNS does not specify which day of the week the food diary record refers to. We solve this problem by exploiting the unique feature of the 1975–1994 NNS, that each participating household is requested to choose *three consecutive days* excluding Sundays and holidays and to report the details of all meals each member had in each day. This means that each household's survey period covers at least two consecutive weekdays, and we decide that school lunch is available to a child if school lunch is reported for at least one of the three days.

Individual reports might not accurately reflect the municipal school lunch status, even for weekdays, for two reasons. First, some children might miss school lunch due to sickness or extra-curricular activities such as excursions. Second, because the NNS does not contain information on the school children attend, school lunch status can be mismeasured if children attend municipal schools outside of the municipality of residence or choose to attend private or national schools, few of which provide school lunch. We presume that the first type of measurement error is small because we use the food diary of three consecutive days and because the NNS instructs households to choose a survey period when the household members have a normal daily life without special events. Measurement error due to children attending municipal schools outside of their municipality is also expected to be small due to the strict Japanese school district system. To address the noise due to children attending non-municipal schools, we exclude prefectures where more than 5 percent of junior-high students attend non-municipal schools in 1994 (shown as shaded areas in Figure 2).

Our sample still exhibits occasional disagreement within a district, in which case we determine the school lunch status based on the majority rule: whether a half or more of junior-high students report having school lunch. To improve accuracy, we exclude districts with only one junior-high respondent and districts with exactly two conflicting answers. School lunch provision at a municipal elementary school is determined analogously. In addition, when the school lunch provision status is obvious from official statistics, we use that information instead of survey response. Details on the procedure to determine the school lunch status are provided in Appendix 3.

To illustrate the accuracy of the imputed school lunch status Table 1 presents the district-level distribution of the percentage of students who report having school lunch. In 2,213 districts (97.4%), the majority of elementary student respondents report having school lunch, consistent with the nearly 100 percent school lunch coverage rate in official reports (Appendix 1). For junior-high students, the reported percentage equals 100% in 65.7 percent and 0% in 20.6 percent of districts, and within-district reporting discrepancy arises in only 13.7 percent districts. Moreover, in the vast majority of districts with mixed reports, the percentage of positive reports is 50 or more, suggesting that the major cause of reporting discrepancies is occasional nonattendance. In Appendix 3, we also confirm that the derived school lunch status is highly consistent with the official statistics. Our final sample consists of 8,477 junior-high students and 9,794 elementary students from 2,265 districts, of which 498 districts do not provide school lunch in their municipal junior highs.

4.3. Weight measures, control variables, and summary statistics

Our outcome variables include BMI, binary indicators for obesity and underweight, and another measure of weight-for-height widely used in Japan called Percentage of overweight (POW). POW is the proportion of weight deviation to the standard weight-for-height by age and gender, defined as ([weight] – [standard weight]) / [standard weight]. We take standard weight-for-height for Japanese children from Murata and Ito (2003). We also examine height in some analysis. For BMI and height, we use both the raw value and the *z*-score that is normalized by gender, age, and five-year cohort.

Defining obesity and underweight requires caution because the distribution of children's BMI varies substantially by age, gender, and race. We use the extended International Obesity Task Force (IOTF) definition (Cole and Lobstein, 2012) and the POW definition. The IOTF definition is the BMI-based definition widely used to account for racial differences in growth patterns. The cutoff of being obese for each age-gender group is determined based on the percentage rank corresponding to BMI 30.0 at age 17.5. We use BMI 25.0 for the obesity cutoff instead of the widely-used 30.0, following the recommendation by the Japan Society for the Study of Obesity (Kanazawa et al., 2002). Underweight is defined analogously: the percentage rank corresponding to BMI 18.5 at age 17.5 is used to determine the underweight cutoff. These IOTF cutoffs minimize across-age variation in the prevalence of obesity and underweight and thus fit well in our DID framework. For Asian children, the IOTF definition is preferred over the WHO reference, which is based on American children's BMI distribution (de Wilde et al., 2013). We use the gender- and age-specific BMI distributions for Japanese children reported in Kato et al. (2011).

Under the POW definition, children whose weight exceeds their standard weight by 20% (i.e., POW>20%) are categorized as obese. Likewise, POW less than −20% is underweight. The POW definition has an advantage over BMI-based measures as BMI increases with height for children in puberty (Sugiura and Murata, 2011).

Table 2 shows the summary statistics of height and weight measures of elementary and junior-high students by treatment status for the full sample (Panel (a)) and the two low-SES subsamples (Panels (b) and (c)). Junior-high students have larger BMIs than elementary students, but the other weight measures show no clear age trends, as expected. The two obesity measures are reasonably consistent. The two underweight measures are different in level but similarly lack strong age trends. Children in the two socioeconomically disadvantaged subsamples are consistently shorter than children in the full sample, but the weight measures show no clear difference. In all three samples, comparison between the treatment and control groups among elementary students shows little difference, except that BMI, POW, and obesity prevalence are smaller in the treatment districts than in the control districts. In contrast, among junior high students they are all larger in the treatment districts than in the control districts in all three samples, consistent with the formal DID results below.

To make our DID estimator more credible, we account for differential age-trends specific to each area, which cannot be addressed by district fixed effects, by including controls for various child, parental, and household characteristics. For child characteristics we use gender-specific age dummies, with separate age dummies for 12-year-old elementary students and 12-year-old junior-high students. Parental characteristics include age, height, BMI, and occupation dummies of the father and mother.[12] Household characteristics include the coresidence of father, grandfather, and grandmother, the number of children in the household, and the percentage ranking of monthly household expenditure per household member.

Table 3 reports the summary statistics of the control variables by treatment status with the normalized difference between the two groups.[13] The table shows that junior-high school lunch is more likely to be provided in districts where children live with their grandparents, parents work in agriculture/fisheries/forestry, and per-member expenditure is lower. The absolute normalized differences in the control variables between the treatment and control groups are overall small and never exceed the threshold value of 0.25, where values above 0.25 are considered problematic (Rubin, 2001). IPTW and propensity-score trimming make these differences even smaller, as shown in Appendix 2 (Table A1).

4.4. Determinants of school lunch provision

---

[12] Because the NNS does not record kinship among household members, we regard 27- to 59-year-olds in a child's household as parents and those 60 years old or older as grandparents. Mean values are used in the rare case of multiple "fathers" or "mothers."

[13] Many DID studies conduct *t*-tests for differences in covariate means between treatment and control groups. However, the normalized differences are more effective for assessing the difficulty of adjusting for covariate differences between the groups than the *t*-statistics (Imbens, 2015).

Tables 2 and 3 indicate that junior-high lunch status is not purely random and that a mere comparison of outcome variables between junior-high students with and without school lunch might be biased. Even if we control for observable district characteristics, the unconfoundedness assumption might be violated. Our DID estimation addresses this non-randomness and yields unbiased causal estimates under the common trend assumption. Specifically, the assumption requires that the school lunch provision be determined orthogonally to children's growth in height and weight from age 9 to 15 (after controlling for district characteristics that vary over age).

To assess this premise, it is useful to understand the determinants of school lunch provision at junior highs. We first conduct a district-level Logit regression analysis in which we regress the *NoSchoolLunch* dummy on mean height, mean BMI, the prevalence rates of obesity and underweight among children 1- to 11-year-old, and other various district characteristics. We use districts with five or more respondents of age 1 to 11. To minimize the influence of age-composition among children, we use mean *z*-scores of height and BMI that are normalized by gender, age, and five-year cohort, and use the IOTF definitions for obesity and underweight. We include residents' characteristics defined as aggregated NNS values: the number of NNS participants, age composition, median percentage ranking of per-member household expenditure, mean household size, occupational composition among 23- to 54-year-old workers, and the proportion of working women. Also included are prefectural population density obtained from Statistics Bureau, Ministry of Internal Affairs and Communications (2012), dummies for municipal size, 47 prefecture specific effects or six region block specific effects, and year dummies. Their summary statistics are provided in Appendix 4 (Table A3).

Table 4 reports the regression results. Model 1 controls for region block fixed effects and Model 2 controls for prefecture fixed effects. Junior-high school lunch is less common in larger municipalities. The coefficient on the prefectural population density is significantly negative only in the prefecture fixed-effect model, implying a higher likelihood of school lunch provision in prefectures with a growing population. Consistent with this, school lunch is less common in “older” areas with a larger fraction of the elderly population. The estimated coefficients of year dummies (reported in Appendix 4, Table A4) indicate a significant increase in school lunch provision over time, consistent with both the nationwide and sample trends.

On the other hand, none of the other variables, including those directly related to obesogenic environments, are significant. Most importantly, in both specifications, mean height, mean BMI, obesity rate, and underweight rate among elementary and preschool children are independently and jointly nonsignificant, which suggests that the municipal school lunch provision is unrelated to the stature, obesity, and underweight of children. The lack of significance of the expenditure level implies that the municipal decision on school lunch is not driven by the economic deprivation of residents. Occupation is a major determinant of obesogenic environments (Maruyama and Nakamura, 2018), but occupational composition variables are not significant. These findings suggest the absence of a direct link between the school lunch provision and obesogenic environments around children, supporting our common trend assumption.[14] To further address concerns about the possible

[14] We also conduct a substantial archival work, examining media coverages, research articles, and minutes of municipal assemblies, to understand the determinants of school lunch provision, but little information has been identified. Recently the coverage rate of junior-high school lunch started to rise (Figure A1), and one reason is the rising concerns and awareness over child nutrition (Mainichi, 2016). In the Third Basic Program for Shokuiku Promotion of 2016, the government introduced an agenda to increase school lunch provision at municipal junior highs to promote healthy eating. Reverse causality from local obesity prevalence to municipal school lunch provision is of little concern during our study period 1975–1994, however, because until the 2000s few were aware of child obesity issues in Japan even among health professionals (Yoshinaga, 2012). The lack of significance of the variables on children’s weight is consistent with this view.

violation of the common trend assumption, we conduct a number of robustness tests, as discussed in the next section.

5. Results

5.1. Main results

Figure 4 plots year-adjusted means of height and BMI by gender in the full sample and the low-SES subsamples over ages 9 to 15 for the treatment and control groups. The vertical red line indicates age 12, the threshold age between elementary and junior-high students. In this figure, 15-year-olds include both current- and post-junior high students. For all three samples, the trends in mean height are highly similar between the treatment and control groups for both genders, highlighting little difference in growth patterns. The differences in BMI trends are also small in the full sample. Among children with non-white-collar fathers, while elementary boys' BMI is similar between the groups, elementary girls' BMI is lower for the treatment than the control groups, but BMI of the treatment group exceeds that of the control group at ages 13 and 14 for both genders. Among children with low household expenditure, fluctuations in BMI are large, but the overall trends are similar to those for children with non-white-collar fathers.

Table 5 reports the estimated effects of no school lunch from (a) junior-high OLS, as specified in Eq. (1), (b) DID regression, and (c) DID regression with IPTW and propensity-score trimming, in Panels (a)–(c) respectively.[15] Whereas we find no statistically significant effect on any outcome measure across all three specifications for the full sample, the subsample results reveal fairly robust evidence that school lunch reduces BMI, POW, and obesity for children from lower socioeconomic backgrounds. In the subsample of children with fathers in non-white-collar occupations, the estimated effects are significantly positive on BMI, BMI *z*-score, POW, and both of the obesity

[15] Full regression results are available from the authors upon request.

measures in all three specifications, at least at the 10 percent level. The estimates from the base DID imply that the lack of school lunch increases BMI by 0.403 and increases obesity prevalence by about 3 percent under both IOTF and POW definitions, the estimates from the DID with IPTW and trimming being even larger. In the subsample of children with low household expenditure, while none of the estimated effects from the OLS are significant except for the obesity status under POW definition, those from the base DID and DID with IPTW and trimming are all significantly positive on all outcomes except for the underweight measures. The base DID estimates imply that the lack of school lunch increases BMI by 0.418 and increases obesity prevalence by about 5 percent under both definitions, with the estimates from the DID with IPTW and trimming being larger again. We find no evidence that school lunch affects underweight prevalence in any specification regardless of the sample used.[16]

5.2. Quantile DID

To explore effect heterogeneity over the baseline distribution of height and weight, we conduct raw quantile DID (QDID) analysis, in which DID estimates are obtained by taking the difference of two differences for each of 19 equidistant intervals between the 5th and 95th percentiles of the dependent variable.[17] Figure 5 shows the QDID estimates for height in z-score, BMI in z-score, and POW. For all three samples, the estimates for height are close to zero across quantiles. In contrast, the estimates

[16] We also conduct a subsample analysis of children whose per-member household expenditure is *above* the median and children with fathers in white-collar occupations or without fathers, and none of the estimated effects are significant.

[17] We refrain from formal quantile regression analysis because the econometric theory for the quantile treatment effect in DID settings with control variables is under development and because our data and setting are not ideal for such analysis.

for BMI and POW are mostly positive and larger for higher quantiles for all three samples, implying the variance-reduction effect of school lunch for weight. This pattern is particularly large for the highest quantiles of the low-SES subsamples. Obese children are the driving force behind the significant weight-reduction effect of school lunch among the low-SES subsamples.

6. The Robustness of Results

6.1. Permutation test

To assess the potential underestimation of standard errors due to correlation of the error terms among children close in age and location, we implement permutation tests in the spirit of Bertrand et al. (2004) and Abadie et al. (2010). In this test, the treatment *NoSchoolLunch* is randomly assigned to control districts holding the same frequency as in the original sample, and the base DID model is estimated based on the hypothetical treatment assignment. Repeating this procedure 1,000 times yields the distribution of the estimated placebo treatment effect. As detailed in Appendix 5 (Figure A3), the implied statistical significance in this test is highly consistent with the corresponding results in Table 5, suggesting that the statistical significance of our estimated school lunch effects is not due to the misspecification of correlation structure.

6.2. Falsification test: regression analysis of height

Our DID framework relies on the common trend assumption that the school lunch provision is determined orthogonally to children’s growth in height and weight from age 9 to 15 (after controlling for covariates that vary over age). This assumption might be violated if there are systematic differences in the growth pattern between the treatment and control districts, especially in the timing of puberty onset. The use of POW, POW-based obesity/underweight definitions, IPTW, and propensity-score trimming should mitigate potential bias. We further conduct a falsification test in which we use height instead of the weight measures as a regressand. Because height is determined primarily by genetic factors and early-life environment (Beard and Blaser, 2002), school lunch should not have a strong, immediate effect on height. Thus, the coefficient of the interaction of *JuniorHigh dummy* and *NoSchoolLunch dummy* would reflect differences in children's pre-determined growth pattern between the treatment and control districts. As shown in Table 6, none of the DID estimates are significant for the full sample and the low-SES subsamples, providing additional support for the common trend assumption.

6.3. Other robustness checks

Further robustness checks support the robustness of our results, as detailed in Appendix 5. Our results might be confounded by the differential effects of attending a junior high by urbanicity and survey timing. While the school lunch status is significantly associated with the municipality size and survey year (Table 4), the transition from elementary to junior-high schools might involve lifestyle alterations, such as an increase in time spent for studying or school sports club activities, whose effect on weight might vary with urbanicity and change over time. To address this concern, we add interactions of *JuniorHigh* dummy with five-year period dummies, prefectural population density, and municipality size variables to the DID regression model. Furthermore, we also test specifications with triple interactions of the *JuniorHigh* dummy, year dummies, and prefecture dummies to the DID regression model to allow for differential effects of attending junior high by prefecture-year. Our results are robust to these modifications.[18]

[18] Another robustness test of the common trend assumption would be to include linear age trends specific to each district in the DID analysis. However, this is not appropriate in our setting because the growth curves of height and weight over age are highly nonlinear. Additionally, it is not feasible because the number of children is too small in many districts, especially in the subsamples.

We implement a synthetic control method proposed by Abadie et al. (2010) to compare the outcome trends of the treatment group with an artificial control group with more similar pre-treatment trends than the original control group. The results confirm the robustness of our findings.

Further specification tests are conducted. First, we estimate the DID regression model using a less restrictive criterion for the selection of prefectures: we include in our sample prefectures with at least 90 percent of junior-high students attending municipal schools as of 1994 instead of 95 percent, despite a potential concern about bias due to students attending non-municipal schools. This modification does not change our findings substantially. Second, we estimate the same model excluding 12- and 15-year-olds. Among 12-year-olds, some are elementary and others are junior-high students, depending on the birth month. Similarly, some of the 15-year-olds are current junior-high students and others have finished junior high. Hence, our results might be affected by potential differential effects of the birth month on children's height and weight between the treatment and control groups. We confirm the robustness of our findings to this exclusion and also find highly similar effects of the birth month on height for both groups.

6.4. Misclassification in school lunch status

The imputed school lunch provision status for junior-high schools might be subject to measurement error, potentially causing estimated effects attenuated toward zero. This might be an explanation for the insignificant effects of school lunch on underweight. If that is the case, however, the estimated positive effects of the lack of school lunch on BMI and obesity can be considered the lower bound of the true effect.

7. Long-Term Effects

Does the weight reduction effect of school lunch for low-SES children last after they graduate from junior high and stop having school lunch? The long-term effect of school lunch is a parameter of policy interest because it can drastically alter the cost-benefit of such programs. To estimate the aftereffect of school lunch, we add 15- to 17-year-old post-junior-high students in the sample and estimate the following model:

$$Y_{id} = X_{id}\beta + \gamma NoSchoolLunch_d + \theta^{DDS} JuniorHigh_{id} \times NoSchoolLunch_d$$
$$+\theta^{DDL} PostJH_{id} \times NoSchoolLunch_d + \mu_d + \epsilon_{id}, \tag{3}$$

where $PostJH_{id}$ is an indicator for post-junior-high students, and θ^{DDS} and θ^{DDL} denote the DID estimators of short- and long-run effects of no school lunch, respectively. The 15- to 17-year-olds in the control group are assumed to have had school lunch for the three years before graduation. This would be a reasonable approximation given the slow expansion in the school lunch coverage during our study period.[19]

Table 7 reports the two estimated coefficients on the *NoSchoolLunch* dummy interacted with the junior-high and post-junior-high dummies. As expected, the estimated effects for junior-high students (henceforth "ongoing effects") are all positive and close to the corresponding results in Table 5. They are all nonsignificant for the full sample, and all significantly positive for the low-SES subsamples, at least at the 10 percent level. The estimated effects of post-junior-high students (henceforth "post-effects") are also all positive and fairly close to the ongoing effect in both significance and magnitude. At the 10 percent significance level, whereas only two out of ten estimates of the post-effects are significant for the full sample, six out of ten are significant for children with non-white-collar fathers, and seven out of ten are significant for children with low household expenditure. For the low-SES subsamples the estimated post-effects are either close to or

[19] Even when a non-negligible fraction of the 15- to 17-year-olds in the control group was exposed to school lunch for less than three years, our estimate of the longer-term effect can be interpreted as a lower bound of the true effect. We refrain from estimating longer-term effects because relocation of individuals is very common after the age of 18.

smaller than the ongoing effects, except for a few estimates from DID with IPTW and trimming, which exceed the ongoing effects.

We repeat the QDID analysis, treating post-junior-high students, instead of current junior high students, as post-treatment observations. The results shown in Figure 6 are highly consistent with those from the main analysis (Figure 5). Although the estimates are overall smaller, the estimates for the BMI and POW in the low-SES subsamples are positive for the highest quantiles.

These findings provide suggestive evidence that the obesity-reducing effect of the school lunch program for children with low socioeconomic backgrounds lasts at least several years after they graduate, implying the role of school lunch in preference and habit formation for low SES children, consistent with the Japanese government's policy to promote healthy eating habits and food knowledge through school lunch.

8. Conclusion and Discussion

We examine the causal effect of a Japanese school lunch program on the weight of junior-high students using data drawn from the 1975–1994 NNS. To account for possible endogeneity of the municipal provision of school lunch, we employ a DID framework that compares differences between elementary and junior-high students across districts with and without school lunch at junior-high schools. Although we find no significant effects when the full sample is used, we find that the lack of school lunch significantly increases the BMI and obesity rate of children from low socioeconomic backgrounds. We find little evidence that school lunch affects underweight prevalence, implying the variance-reduction effect of school lunch. The obesity-reducing effect of school lunch for low SES children appears to last at least several years after graduation, suggesting that the school lunch effect involves not only direct changes in caloric and nutritional contents but also changes in behavior. Our findings are robust to propensity score-based trimming and re-weighting (IPTW), permutation tests, various falsification tests of the common trend assumption, and several specification tests. We also present that municipal school lunch provision for junior-high schools is unrelated to the growth pattern of elementary and preschool children in the area.

This study has the following policy implications. First, school lunch with strict nutritional requirements could reduce obesity. This finding is consistent with the literature in that, while previous findings on American NSLP and SBP are mixed, the majority of studies find obesity reduction effects of SBP, which satisfy stricter nutritional requirements than NSLP (Gleason and Dodd, 2009; Millimet et al., 2010; Millimet and Tchernis, 2013). Second, the compulsory nature of the Japanese school lunch program, with its primary purpose of abolishing stigma, might play a key role in obesity reduction because children with a strong preference for fattening food might avoid school lunch if a choice is allowed. Third, school lunch might have long-term obesity-reducing effects via preference and habit formation, consistent with the educational goals of the Japanese school lunch program of promoting good eating habits and food knowledge (NIEPR 2013). Fourth, the obesity reduction effects of school lunch for low-SES children might have contributed to the absence of significant income gradient in child obesity among elementary-school students (Kachi et al., 2015) and the small income gradient in child health in Japan (Nakamura, 2014).

We find beneficial effects of school lunch only for low SES children. From the efficiency point of view, this finding seems to justify means-testing. However, the compulsory nature of the Japanese program helps to avoid inefficiency due to self-selection and social stigma. The setting where everyone eats together also facilitates food education and potentially brings beneficial peer effects. Means-testing and allowing for choice need to be evaluated in further research.

Data limitation prevents us from using more recent data for this study, and future studies would require the accumulation of microdata of children, with accurate information on both body measurements and school lunch participation. Additionally, previous studies find a positive impact of school lunch quality on academic performance in the UK (Belot and James, 2011) and US (Anderson et al., 2017), and future research includes examining the effect of school lunch on other outcomes in Japan, including cognitive ability, which is infeasible with our data.

Acknowledgments: The authors are grateful for a Grant-in-Aid for Young Scientists (B) No. 25780187 from the Japan Society for the Promotion of Science. We also wish to thank Mayu Fujii, Daiji Kawaguchi, Ayako Kondo, Jui-fen Rachel Lu, Eiji Mangyo, Hideo Owan, Mike Shields, Ryuichi Tanaka, and seminar participants at Hitotsubashi, Nagoya, Osaka, Tokyo, and Toyama Universities for their helpful comments. Both authors declare that they have no conflicts of interest. Official permission to utilize the National Nutrition Survey and the National Health and Nutrition Survey was obtained from the Statistics and Information Department of the Ministry of Health, Labour and Welfare on October 2, 2015 (Permission Number 1002-6) and November 20, 2015 (Permission Number 1120-1), respectively.

References

Abadie A, Diamond A, Hainmueller J, 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105: 493–505.

Angrist J, Pischke, JS, 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Andersen SS, Baarts C, Holm L, 2017. Contrasting approaches to food education and school meals. *Food, Culture & Society* 20 (4): 609–629.

Anderson ML, Gallagher J, Ritchie ER, 2017. School lunch quality and academic performance. *NBER Working Paper* No.23218.

Beard AS, Blaser MJ, 2002. The ecology of height: The effect of microbial transmission on human height. *Perspectives in Biology and Medicine* 45 (4): 475–498.

Baidal JA, Taveras EM, 2014. Protecting progress against childhood obesity: The National School Lunch Program. *New England Journal of Medicine* 371 (20): 862‑1865.

Belot M, James J, 2011. Healthy school meals and educational outcomes. *Journal of Health Economics* 30 (3): 489–504.

Bertrand M, Duflo E, Mullainathan S, 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119 (1): 249–275.

Capogrossi K, You W, 2017. The influence of school nutrition programs on the weight of low-income children: A treatment effect analysis. *Health Economics* 26 (8): 980–1000.

Cole TJ, Lobstein T, 2012. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity* 7 (4): 284–294.

Darmon N, Drewnowski A, 2008. Does social class predict diet quality? *American Journal of Clinical Nutrition* 87 (5): 1107–17.

de Wilde JA, van Dommelen P, Middelkoop BJ, 2013. Appropriate body mass index cut-offs to determine thinness, overweight and obesity in South Asian children in the Netherlands. *PLoS One* 8 (12): e82822. doi:10.1371/journal.pone.0082822.

Evans CEL, Harper CE, 2009. A history and review of school meal standards in the UK. *Journal of Human Nutrition and Dietetics* 22 (2): 89–99.

Fujisawa H, 2008. Disputes of school lunch issues. *Osaka Keidai ronshu* 59 (2): 199–214 (in Japanese).

Gleason PM, Dodd AH, 2009. School breakfast program but not school lunch program participation is associated with lower body mass index. *Journal of the American Dietetic Association* 109 (2): S118–S128.

Gorber CS, Tremblay M, Moher D, Gorber B, 2007. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: A systematic review. *Obesity Reviews* 8 (4): 307–326.

Guariguata L, Whiting DR, Hambleton I, et al., 2014. IDF diabetes atlas: Global estimates of diabetes prevalence for 2013 and projections for 2035, *Diabetes Research and Clinical Practice* 103: 137–149.

Gundersen C, Kreider B, Pepper J, 2002. The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics* 166: 79–91.

Harper C, Wood L, Mitchell C, 2008. *The Provision of School Food in 18 Countries. School Food Trust*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.654.9233&rep=rep1&type=pdf

Hazano S, Nakanishi A, Nozue M, Ishida H, Yamamoto T, Abe A, Murayama N, 2017. The relationship between household income and food intake of Japanese schoolchildren. *The Japanese Journal of Nutrition and Dietetics* 75 (1): 19–28.

Hernandez DC, Francis LA, Doyle EA, 2011. National School Lunch Program participation and sex differences in Body Mass Index trajectories of children from low-income families. *Archives of Pediatric Adolescent Medicine* 165 (4): 346–353.

Hinrichs P, 2010. The effects of the National School Lunch Program on education and health. *Journal of Policy Analysis and Management* 29 (3): 479–505.

Hofferth SL, Curtin S, 2005. Poverty, food programs, and childhood obesity. *Journal of Policy Analysis and Management* 24 (4): 703–726.

Imbens GW, 2015. Matching methods in practice: Three examples. *Journal of Human Resources* 50 (2): 373–419.

Jo Y, 2014. Research on the food and agricultural education in South Korea: Comparative study of school lunch programs and dairy educational farms in South Korea and Japan. *Nihon University Ph.D. Thesis* (in Japanese).

Jomaa LH, McDonnell E, Probart C, 2011. School feeding programs in developing countries: Impacts on children's health and educational outcomes. *Nutrition Reviews* 69 (2): 83–98.

Kachi Y, Otsuka T, Kawada T, 2015. Socioeconomic status and overweight: A population-based cross-sectional study of Japanese children and adolescents. *Journal of Epidemiology* 25 (7): 463–469.

Kanazawa M, Yoshiike N, Osaka T, 2002. Criteria and classification of obesity in Japan and Asia-Oceania. *Asia Pacific Journal of Clinical Nutrition* 11(Suppl): S732–S737.

Katanoda K, Nitta H, Hayashi K, Matsumura Y, 2005. Is the national nutrition survey in Japan representative of the entire Japanese population? *Nutrition* 21 (9): 964–966.

Kato N, Takimoto H, Sudo N, 2011. The cubic functions for spline smoothed L, S and M values for BMI reference data of Japanese children. *Clinical Pediatric Endocrinology* 20 (2): 47–49.

Kawaraya C, Mori H, 2009. The dietary habits of junior high school students: Eating vs. not eating school lunches. *Osaka Shoin Women's University Faculty of Liberal Arts collected essays* 46: 77–90.

Kingston T, 2019, July 31. Focaccia off! Italy bans school packed lunches. *The Times.* Retrieved from ProQuest.

Kodama H, 2010. Problems of underweight in young females and pregnant women in Japan. *Japan Medical Association Journal* 53 (5): 285–289.

Lake A, Townshend T, 2006. Obesogenic environments: Exploring the built and food environments. *Journal of the Royal Society for the Promotion of Health* 126 (6): 262–267.

Lee Q, 2013. Childhood socioeconomic status and obesity in Japan. *Quarterly of Social Security Research* 49 (2): 217–229 (in Japanese).

Li J, Hooker NH, 2010. Childhood obesity and schools: Evidence from the national survey of children's health. *Journal of School Health* 80 (2): 96–103.

Mainichi, 2016, July 9. *Yokosuka City to Introduce Complete School Lunch at Junior High Schools*. Retrieved from mainichi.jp (in Japanese).

Maruyama S, Nakamura S, 2015. The decline in BMI among Japanese women after World War II. *Economics & Human Biology* 18: 125–138.

Maruyama S, Nakamura S, 2018. Why are women slimmer than men in developed countries? *Economics & Human Biology* 30: 1–13.

MEXT, 2015. *Results of Current Status Survey on School Meals* (in Japanese). http://www.mext.go.jp/b_menu/houdou/27/02/1354919.htm

Meyer BD, Sullivan JX, 2003. Measuring the well-being of the poor using income and consumption. *Journal of Human Resources* 38 (S4): 1180–1220.

MHLW, 2003. *Report on the 2002 National Health and Nutrition Survey* (in Japanese). http://www.mhlw.go.jp/bunya/kenkou/eiyou-chosa2-01/index.html

MHLW, 2016. *Overview of the results from the National Nutrition Survey 2015* (in Japanese). http://www.mhlw.go.jp/file/04-Houdouhappyou-10904750-Kenkoukyoku-Gantaisakukenkouzoushinka/kekkagaiyou.pdf

Millimet DL, Tchernis R, Husain M, 2010. School nutrition programs and the incidence of childhood obesity. *Journal of Human Resources* 45 (3): 640–654.

Millimet DL, Tchernis R, 2013. Estimation of treatment effects without an exclusion restriction: With an application to the analysis of the school breakfast program. *Journal of Applied Econometrics* 28 (6): 982–1017.

Ministry of Education, 1994. *Statistical Abstract of Education* (in Japanese). http://www.mext.go.jp/b_menu/toukei/002/002b/koumoku.html

Mirtcheva DM, Powell LM, 2013. National School Lunch Program participation and child body weight. *Eastern Economic Journal* 39: 328–345.

Miyawaki A, Lee JS, Kobayashi Y, 2019. Impact of the school lunch program on overweight and obesity among junior high school students: a nationwide study in Japan. *Journal of Public Health* 41 (2): 362–370.

Murata M, Itoh K, 2003. On the optimal weight of school age children. In: Murata M (ed.). *Health Labour Sciences Research Grant Report for FY 2002 on Nutrition, Exercise and Recreation from the Viewpoint of Health and QOL in Children* (in Japanese).

Nakamura S, 2014. Parental income and child health in Japan. *Journal of the Japanese and International Economies* 32: 42–55.

National School Health Center of Japan, 1986. *The School Lunch Data Book: Fiscal Year 1986*. Tokyo: Daiichi Hoki Shuppan (in Japanese).

National School Health Center of Japan, 1990. *The School Lunch Data Book: Fiscal Year 1990*. Tokyo: Daiichi Hoki Shuppan (in Japanese).

National School Health Center of Japan, 1991. *The School Lunch Data Book: Fiscal Year 1991*. Tokyo: Daiichi Hoki Shuppan (in Japanese).

Ng M, Fleming T, Robinson M, et al., 2014. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 384 (9945): 766–781.

Nickell S, 1981. Biases in dynamic models with fixed effects *Econometrica* 49 (6): 1417–1426.

NIEPR (National Institute for Educational Policy Research), 2013. *School Lunch Program in Japan*. https://www.nier.go.jp/English/educationjapan/pdf/201303SLP.pdf

Nozue M, 2011. The contribution of school lunch to dietary intakes in school children and application of dietary reference intakes: A case study of fifth-grade school children. *Ph.D. Thesis, Kagawa Nutrition University* (in Japanese).

Nozue M, Jun K, Ishihara Y, 2010. Differences in food consumption and distribution of meals between the days with or without school lunches among 5th grade elementary school students. *Japanese Journal of Nutrition and Dietetics* 68 (5): 298–308 (in Japanese).

Ono N, 2007. Introduction to the review of Article 2 of the School Lunch Implementation Standard. *Journal of Humanities and Social Sciences* 23: 50–61 (in Japanese).

Peterson C, 2014. Investigating the historic long-term population health impact of the US National School Lunch Program. *Public Health Nutrition* 17 (12): 2783–2789.

Reilly JJ, Kelly J, 2011. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: Systematic review. *International Journal of Obesity* 35: 891–898.

Rubin DB, 2001. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology* 2: 169–188.

Rutledge JG, 2015. From charity to security: the emergence of the National School Lunch Program. *History of Education* 44 (2): 187–206.

Sanborn AA, 2017. More than a meal: School lunch in Japan. *Education About ASIA* 22 (1): 45–49.

Schanzenbach DW, 2009. Do school lunches contribute to childhood obesity? *Journal of Human Resources* 44 (3): 684–709.

Shrewsbury V, Wardle J, 2008. Socioeconomic status and adiposity in childhood: A systematic review of cross-sectional studies 1990–2005. *Obesity (Silver Spring)* 16: 275–84.

Smith, JA, Todd PE, (2005). Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators? *Journal of Econometrics* 125: 305–335.

Smith TA, 2017. Do school food programs improve child dietary quality? *American Journal of Agricultural Economics* 99 (2): 339–356.

Statistics Bureau, Ministry of Internal Affairs and Communications, 2012. *Historical Statistics of Japan*. https://www.stat.go.jp/naruhodo/dl/zuhyou/c1s_origin/y680203000.xls

Sudharsanan N, Romano S, Cunningham SA, 2016. School breakfast receipt and obesity among American fifth- and eighth-graders. *Journal of the Academy of Nutrition and Dietetics* 116 (4): 599–607.

Sugiura R, Murata M, 2011. Problems with body mass index as an index to evaluate physical status of children in puberty. *Pediatrics International* 53: 634–642.

Takahashi T, MatsuuraY, Osawa S, et al., 1983. Survey on food intake of elementary and junior high school children in Ibaraki prefecture: With special reference to milk, dairy products, and green and yellow vegetables. *Japanese Journal of Nutrition and Dietetics* 41 (1): 21–29.

Takimoto H, Yoshiike N, Kaneda F, Yoshita K, 2004. Thinness among young Japanese women. *American Journal of Public Health* 94 (9): 1592–1595.

Tickle L, 2014, May 20. Free school meals: £56k's being spent to give children a sandwich. *Guardian*. Retrieved from www.theguardian.com.

Togashi K, Masuda H, Rankinen T, 2002. A 12-year follow-up study of treated obese children in Japan. *International Journal of Obesity* 26: 770–777.

von Hinke Kessler Scholder S, 2013. School meal crowd out in the 1980s. *Journal of Health Economics* 32: 538–545.

Yoshinaga M, 2012. Current status and strategies for obesity and metabolic syndrome in children and adolescents. *Pediatric Cardiology and Cardiac Surgery* 28 (2): 103–109 (in Japanese).

Yoshinaga M, Ichiki T, Tanaka Y, et al., 2010. Prevalence of childhood obesity from 1978 to 2007 in Japan. *Pediatrics International* 52: 213–217.

Figure 1. School lunch scenes in Japan

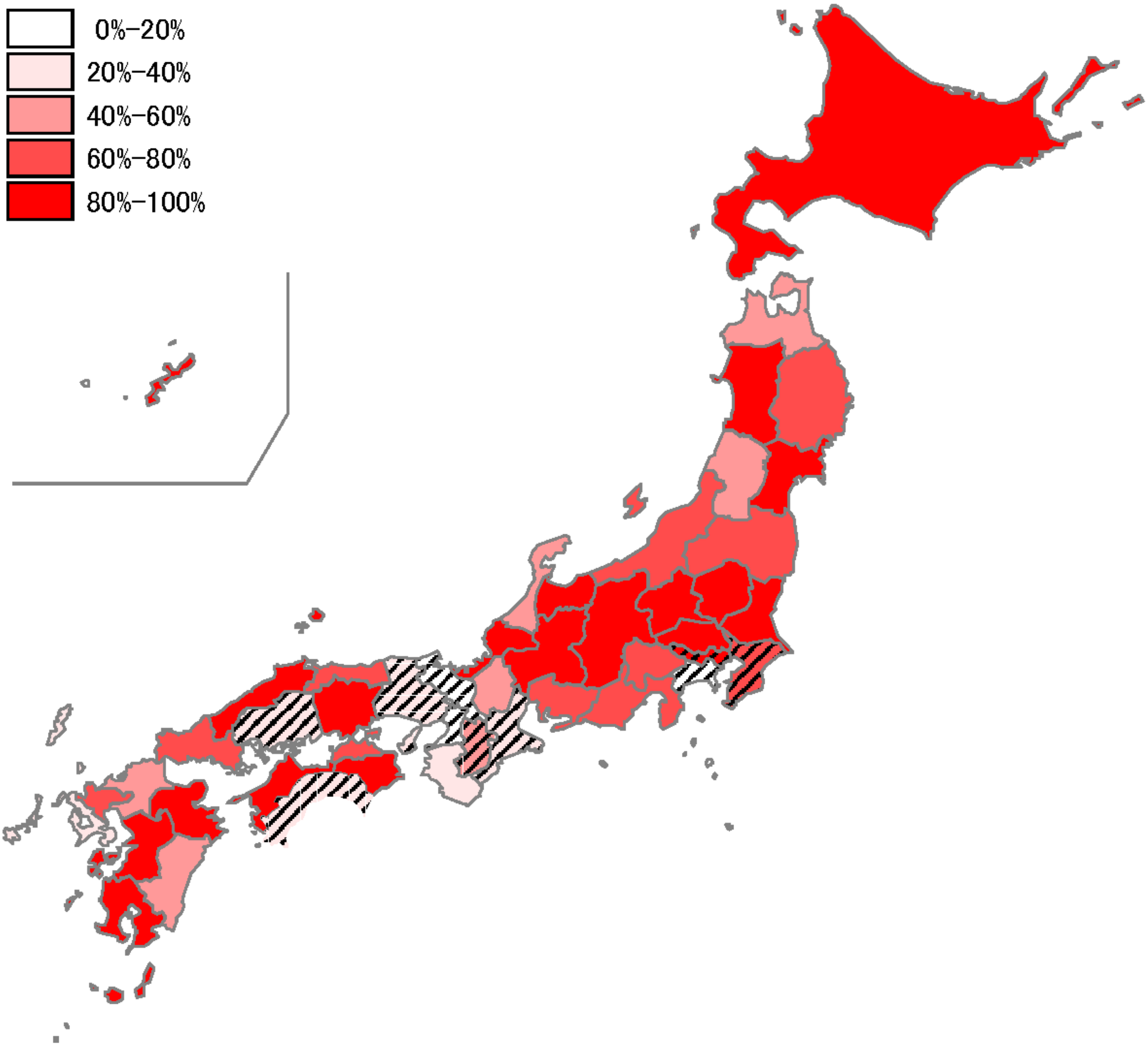
a) Students serving school lunch

Source: “School Lunch at Oyodo Junior High, Osaka City,” March 8, 2017, retrieved from Asahi Shimbun Photo Archive.

b) Students eating together

Source: “School Lunch at Aoyama Junior High, Beppu City,” January 24, 2007, retrieved from Asahi Shimbun Photo Archive.

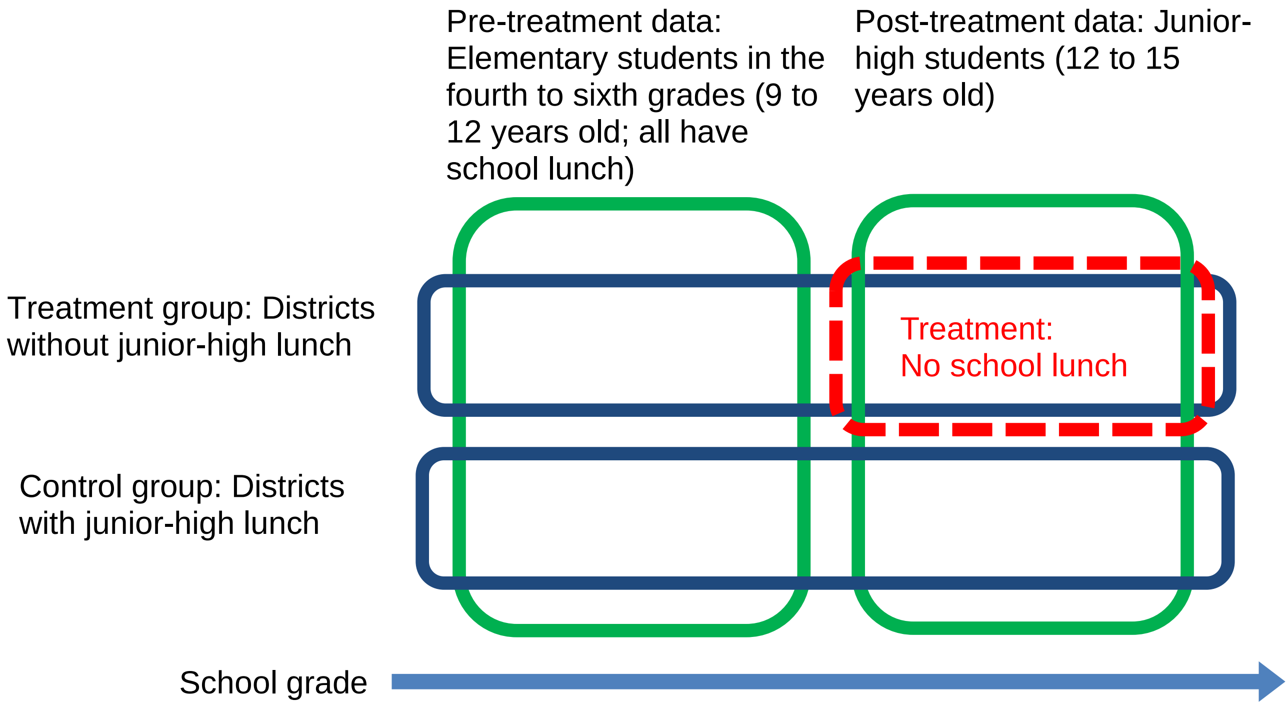
Figure 2. School lunch coverage rate for municipal junior-high students in 1985 by prefecture



Source: National School Health Center of Japan (1986)

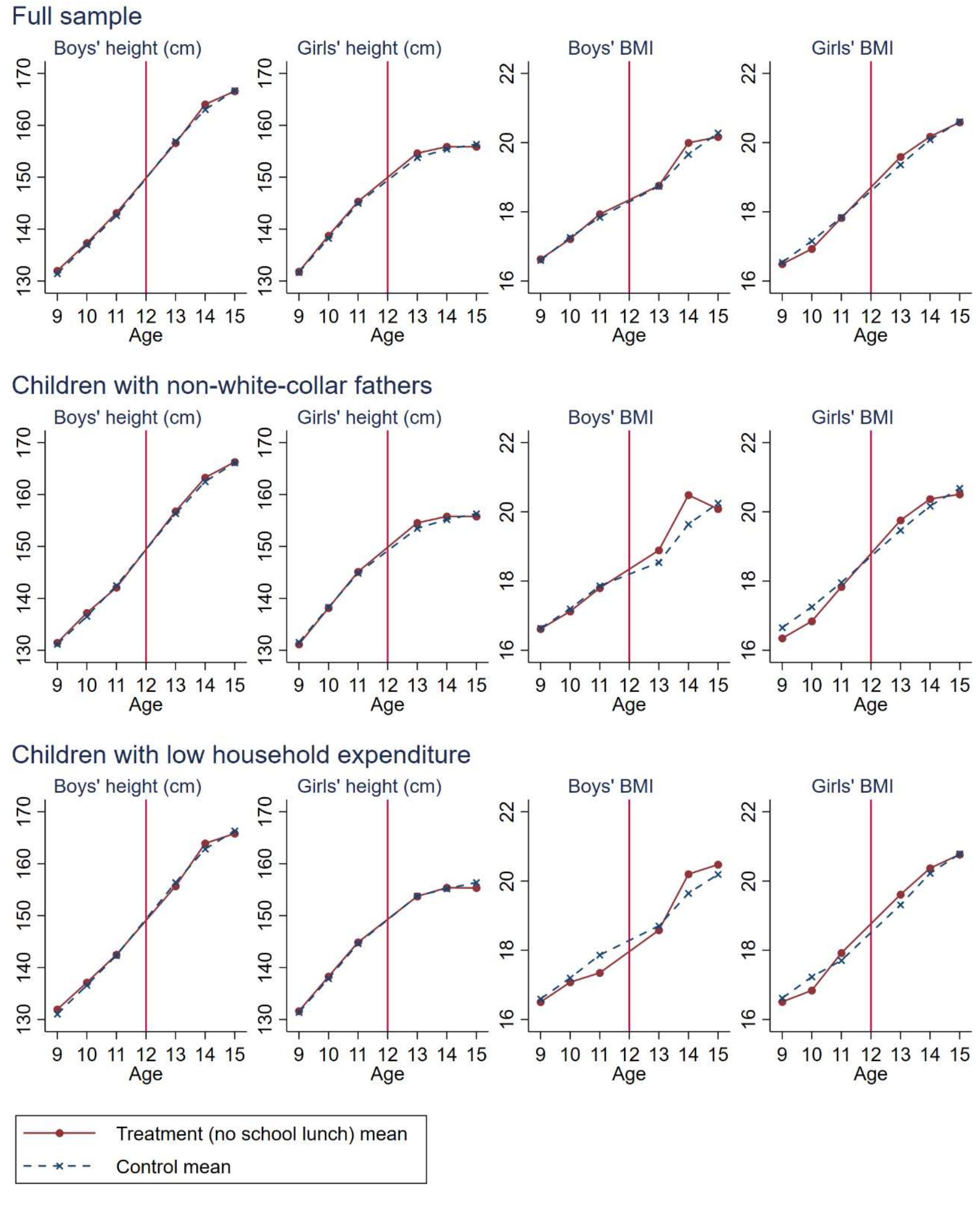
Note: Shaded areas indicate the prefectures excluded from our sample.

Figure 3. Illustration of the DID framework



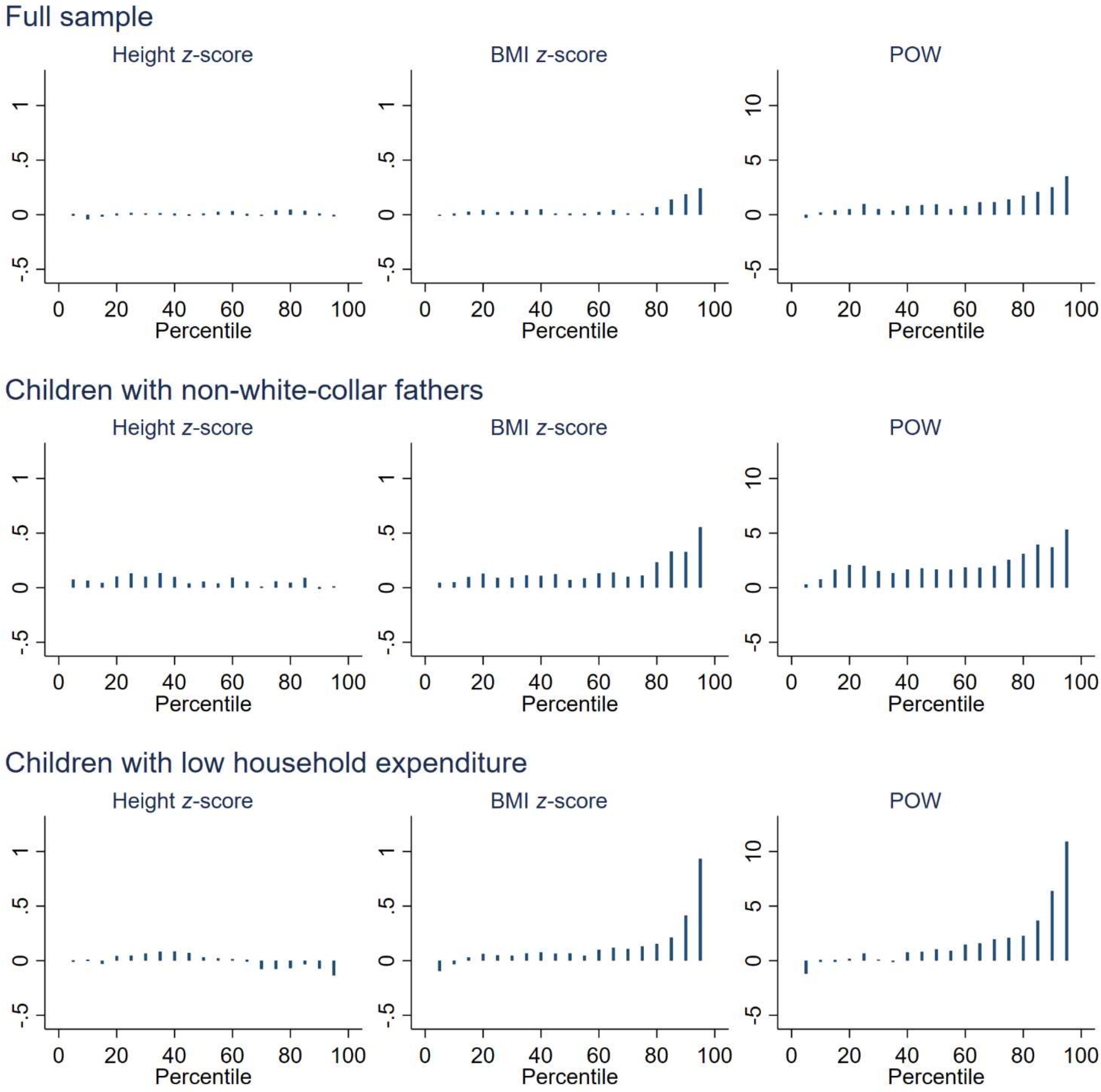
Note: In the Japanese compulsory education system, first to sixth graders attend elementary schools and seventh to ninth graders attend junior highs. A child's school grade is strictly determined by the child's age on April 2nd. Because the NNS is conducted in November each year, fourth to sixth graders in our sample are 9 to 12 years old, and the seventh to ninth graders are 12 to 15 years old.

Figure 4. Means of height and BMI across age



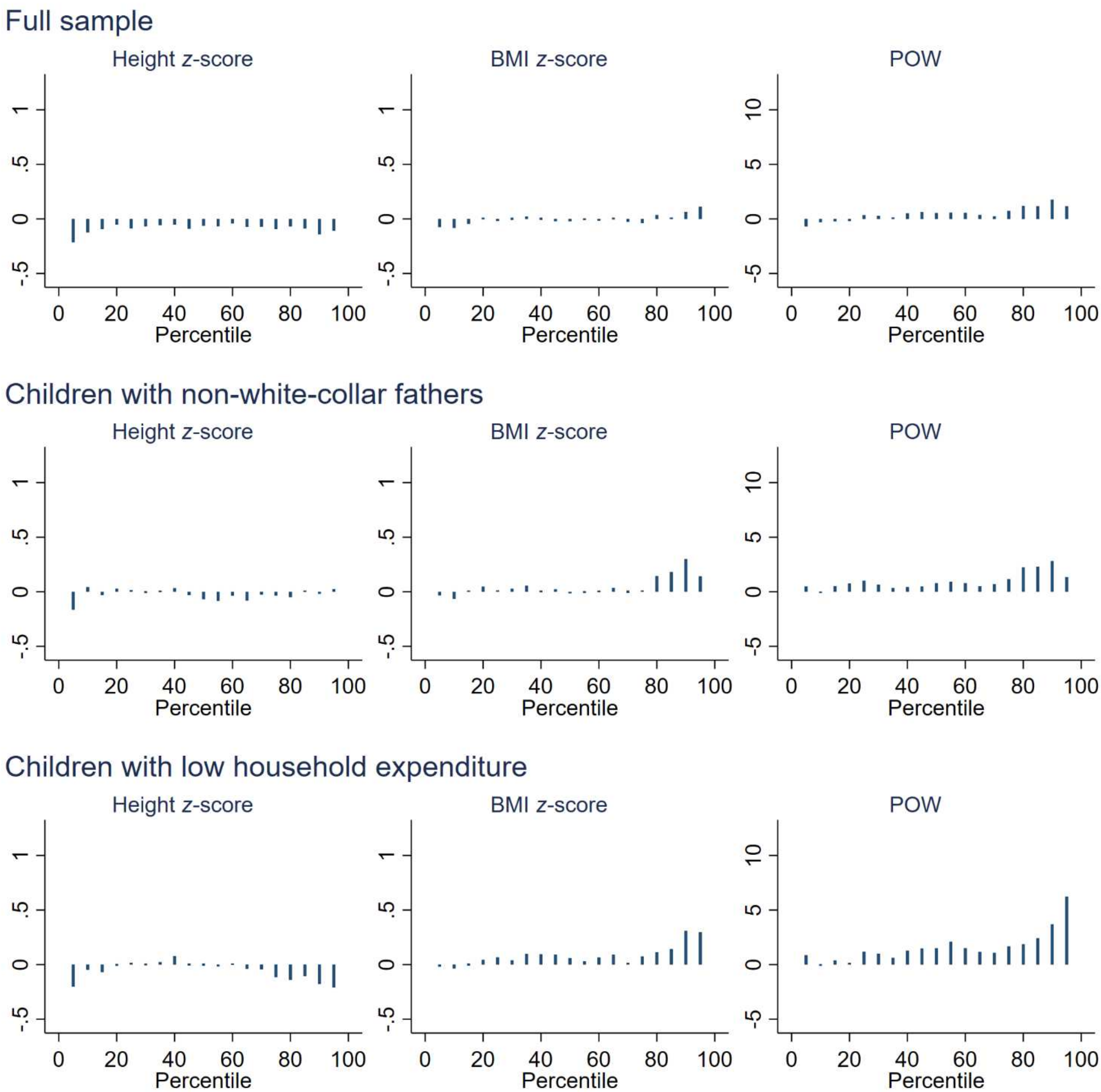
Note: The lines plot year-adjusted means. To adjust for time trends and over-time changes in survey-year composition, each year's value is adjusted by the deviation from the aggregate mean. The vertical red line indicates age 12, the threshold age that divides elementary and junior-high students. While the sample for the main DID analysis only covers elementary and junior-high students, the graph includes 15-year-old post-junior-high students. Values for 12-year-olds are omitted because our sample excludes them for 1975–1985 due to data limitations.

Figure 5. Quantile DID estimates of the effect of no school lunch



Note: The estimates are obtained by calculating the difference of differences with respect to each of the 19 equidistant intervals between the 5th and 95th percentiles.

Figure 6. Quantile DID estimates of the effect of no school lunch for post-junior-high students



Note: The estimates are obtained by calculating the difference of differences with respect to each of the 19 equidistant intervals between the 5th and 95th percentiles. In this Figure, post-junior-high students, instead of current junior-high students, are treated as post-treatment observations.

Table 1. The proportion of students reporting school lunch by district

% of students who reported having school lunch at least once during the survey period	Elementary students		Junior-high students	
No student	36	(1.6%)	468	(20.6%)
more than 0% and less than 50%	23	(1.0%)	60	(2.6%)
50% or more and less than 100%	461	(20.3%)	252	(11.1%)
All students	1,752	(77.1%)	1,492	(65.7%)
Total number of districts	2,272	(100%)	2,272	(100%)

Note: The unit of observation is a district. The figures are based on the sample before we exclude districts with exactly two conflicting reports and districts where 50% or less of elementary students had school lunch (see Appendix 3). Districts with one or no reports are not included.

Table 2. Summary statistics: Outcome variables

	Elementary students				Junior-high students			
	Districts with no junior-high lunch		Control districts		Districts with no junior-high lunch		Control districts	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
(a) Full sample								
Height (cm)	138.54	8.38	138.46	8.55	158.30	7.66	158.06	7.62
Height in *z*-score	0.034	0.962	-0.031	0.984	0.052	0.969	-0.018	0.982
BMI	17.170	2.400	17.304	2.475	19.657	2.744	19.629	2.646
BMI in *z*-score	-0.032	0.930	-0.008	0.957	0.028	1.004	-0.004	0.956
POW	-0.602	12.797	0.285	13.151	-0.068	13.558	-0.360	12.897
Obese (IOTF)	0.069	0.254	0.077	0.267	0.062	0.242	0.059	0.235
Obese (POW)	0.069	0.254	0.079	0.270	0.071	0.257	0.066	0.248
Underweight (IOTF)	0.159	0.365	0.152	0.359	0.159	0.366	0.155	0.361
Underweight (POW)	0.019	0.137	0.020	0.139	0.020	0.139	0.019	0.135
Number of children	2,081		7,713		1,778		6,699	
Number of districts	498		1,767		498		1,767	
(b) Children with non-white-collar fathers								
Height (cm)	137.94	8.29	138.27	8.53	157.82	7.42	157.64	7.63
Height in *z*-score	-0.034	0.953	-0.051	0.990	0.017	0.945	-0.070	0.999
BMI	17.092	2.370	17.354	2.532	19.793	2.801	19.629	2.684
BMI in *z*-score	-0.050	0.907	0.021	0.987	0.090	1.028	0.001	0.972
POW	-0.571	12.948	0.743	13.481	0.693	13.879	-0.250	13.024
Obese (IOTF)	0.067	0.251	0.083	0.275	0.072	0.258	0.056	0.229
Obese (POW)	0.071	0.258	0.084	0.278	0.080	0.271	0.064	0.244
Underweight (IOTF)	0.164	0.371	0.146	0.354	0.148	0.356	0.155	0.362
Underweight (POW)	0.019	0.135	0.021	0.142	0.022	0.147	0.016	0.127
Number of children	1,023		3,901		863		3,439	
Number of districts	349		1,308		349		1,308	
(c) Children with low household expenditure								
Height (cm)	138.29	8.45	138.22	8.62	157.94	7.62	157.82	7.66
Height in *z*-score	-0.049	0.990	-0.083	0.977	-0.018	0.966	-0.052	0.999
BMI	17.055	2.222	17.284	2.493	19.819	2.988	19.666	2.680
BMI in *z*-score	-0.102	0.857	-0.017	0.969	0.084	1.086	0.012	0.963
POW	-0.994	11.862	0.375	13.182	0.687	14.769	-0.076	12.974
Obese (IOTF)	0.053	0.225	0.078	0.268	0.079	0.271	0.058	0.233
Obese (POW)	0.056	0.229	0.078	0.269	0.090	0.287	0.064	0.245
Underweight (IOTF)	0.162	0.369	0.154	0.361	0.156	0.363	0.149	0.356
Underweight (POW)	0.016	0.125	0.020	0.139	0.015	0.122	0.015	0.122
Number of children	826		3,636		730		3,125	
Number of districts	312		1,230		312		1,230	

Table 3. Summary statistics: Control variables

Variable	Districts with no junior-high lunch		Control districts		Normalized difference
	Mean	Std. Dev.	Mean	Std. Dev.	
Male	0.522	0.500	0.515	0.500	0.014
Age	11.764	1.998	11.814	2.010	-0.025
Father's age	38.133	13.071	38.079	12.950	0.004
Father's height (*z* score by age, sex, and 5-year cohort)	0.024	0.851	-0.037	0.844	0.072
Father's BMI (*z* score by age, sex, and 5-year cohort)	-0.017	0.865	0.013	0.846	-0.035
Father's height and BMI missing	0.256	0.436	0.252	0.434	0.009
Father: white-collar worker (the reference category)	0.356	0.478	0.343	0.474	0.027
Father: laborer	0.296	0.455	0.292	0.453	0.011
Father: self-employed	0.191	0.393	0.185	0.387	0.016
Father: agriculture/fisheries/forestry	0.053	0.223	0.078	0.267	-0.104
Father: other occupation (not working)	0.009	0.089	0.009	0.091	0.002
Without father in household	0.095	0.293	0.093	0.291	0.005
Mother's age	39.553	4.319	39.558	4.404	-0.001
Mother's height (*z* score by age, sex, and 5-year cohort)	-0.019	0.981	-0.073	0.974	0.056
Mother's BMI (*z* score by age, sex, and 5-year cohort)	0.000	0.926	0.093	0.997	-0.097
Mother's height and BMI missing	0.036	0.186	0.036	0.187	-0.001
Mother: white-collar worker (the reference category)	0.170	0.372	0.161	0.365	0.024
Mother: laborer	0.237	0.421	0.245	0.427	-0.020
Mother: self-employed	0.136	0.342	0.130	0.334	0.019
Mother: agriculture/fisheries/forestry	0.056	0.227	0.094	0.288	-0.150
Mother: other occupation (not working)	0.402	0.485	0.369	0.478	0.067
Grandfather in household	0.165	0.371	0.204	0.403	-0.101
Grandmother in household	0.272	0.445	0.317	0.465	-0.099
# of children in household (below 18 years old)	2.276	0.813	2.310	0.777	-0.042
Per-member household expenditure ranking (defined between 0.0 and 1.0 where 0.0 means households with highest expenditures in each survey year)	0.607	0.239	0.643	0.241	-0.147
Number of children	3,859		14,412		
Number of districts	498		1,767		

Note: The last column shows the normalized difference between the treatment and control groups. All statistics are based on non-missing observations. When body measurement data (height and BMI) are missing, dummies for the missing values are used in the regression analysis. Because the household expenditure is reported in intervals that vary by survey year, we construct the percentage rank within all households in each survey year for comparability across years. The ranking is recorded in descending order: a larger value implies a smaller expenditure.

Table 4. District-level Logit regression of *NoSchoolLunch* dummy

Explanatory variable	Model 1		Model 2	
District: mean child height (*z* score)	0.201	(0.160)	0.054	(0.181)
District: mean child BMI (*z* score)	-0.071	(0.202)	-0.107	(0.232)
District: child obesity rate (IOTF BMI 25+)	0.409	(0.492)	0.233	(0.528)
District: child underweight rate (IOTF BMI 18.5-)	-0.279	(0.623)	0.024	(0.687)
District: # of participants	-0.001	(0.003)	-0.001	(0.003)
District: proportion of ages 1–19	1.297	(1.409)	1.532	(1.540)
District: proportion of ages 40–59	1.098	(0.963)	1.307	(1.060)
District: proportion of ages 60+	1.932*	(1.121)	2.124*	(1.266)
District: median per-member household expenditure	-0.641	(0.405)	-0.442	(0.460)
District: mean household size	-0.030	(0.136)	-0.124	(0.158)
District: proportion of laborer	0.209	(0.372)	0.090	(0.401)
District: proportion of self-employed	0.092	(0.441)	0.235	(0.487)
District: proportion of agriculture	-0.296	(0.615)	-0.752	(0.670)
District: proportion of working women	0.587	(0.389)	0.185	(0.449)
Prefectural population density (1,000 person/km^2)	0.091	(0.159)	-3.898**	(1.613)
Municipal size: 11 largest cities	2.927***	(0.273)	3.434***	(0.328)
Municipal size: cities with 150k+ population	1.602***	(0.192)	1.685***	(0.212)
Municipal size: cities with 50–150k population	1.375***	(0.200)	1.402***	(0.221)
Municipal size: cities with 50k- population	1.090***	(0.235)	1.109***	(0.256)
Region block dummies	Yes		No	
Prefecture dummies	No		Yes	
Year dummies	Yes		Yes	
R^2	0.153		0.268	
Number of districts	2,117		2,075	

Note: Standard errors are in parentheses. The constant term is included in the model but omitted from the table. The estimated coefficients for year dummies are presented in Appendix 4 (Table A4). The reference category for the municipal size dummies is towns and villages, that for the age composition variables is the proportion of ages 20 to 39, and that for the occupational composition variables is the proportion of white-collar workers. Districts with less than five respondents of ages 1 to11 are excluded from the sample. Model 2 shows fewer observations than Model 1 because prefectures without variation in the *NoSchoolLunch* dummy are omitted. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5. Effects of no school lunch

Panel (a): Junior-high OLS

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	8,470	2,262	0.047	0.020	0.170	0.001	0.003	0.003	0.003
			(0.076)	(0.028)	(0.387)	(0.007)	(0.008)	(0.011)	(0.004)
Children with non-white-collar	4,299	1,654	0.234**	0.089**	1.004*	0.021**	0.020*	-0.014	0.006
fathers			(0.111)	(0.042)	(0.565)	(0.010)	(0.011)	(0.015)	(0.006)
Children with low household	3,853	1,540	0.131	0.059	0.574	0.017	0.025**	0.015	-0.002
expenditure			(0.128)	(0.048)	(0.647)	(0.012)	(0.012)	(0.018)	(0.006)

Panel (b): Base DID

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	18,271	2,265	0.093	0.032	0.767	0.007	0.007	0.002	0.005
			(0.091)	(0.035)	(0.469)	(0.009)	(0.01)	(0.014)	(0.005)
Children with non-white-collar	9,226	1,657	0.403***	0.152***	2.076***	0.031**	0.026*	-0.023	0.010
fathers			(0.135)	(0.052)	(0.707)	(0.013)	(0.014)	(0.020)	(0.008)
Children with low household	8,317	1,542	0.418***	0.167***	2.266***	0.051***	0.052***	-0.004	0.005
expenditure			(0.152)	(0.057)	(0.774)	(0.015)	(0.015)	(0.021)	(0.007)

Panel (c): DID with IPTW and propensity-score trimming

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	13,909	1,728	0.085	0.036	0.763	0.004	0.008	0.002	0.005
			(0.111)	(0.042)	(0.562)	(0.011)	(0.011)	(0.017)	(0.005)
Children with non-white-collar	6,568	1,178	0.515***	0.197***	2.673***	0.038**	0.042**	-0.021	0.013
fathers			(0.175)	(0.066)	(0.898)	(0.017)	(0.018)	(0.023)	(0.008)
Children with low household	6,305	1,175	0.535***	0.223***	2.913***	0.061***	0.056***	-0.016	0.012
expenditure			(0.168)	(0.064)	(0.846)	(0.016)	(0.017)	(0.025)	(0.007)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables in each regression, see Table 3. OLS in Panel (a) also controls for NNS-based district characteristics listed in Table 4. The number of districts in Panel (a) is smaller than that in Panel (b) because of missing district characteristics. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6. The falsification test: regression analysis of height

Panel (a): Base DID

Sample	# children	# districts	Height	Height z-score
Full sample	18,271	2,265	-0.002	0.002
			(0.215)	(0.033)
Children with non-white-collar fathers	9,226	1,657	0.412	0.065
			(0.289)	(0.045)
Children with low household expenditure	8,317	1,542	0.199	0.018
			(0.344)	(0.053)

Panel (b): DID with IPTW and propensity-score trimming

Sample	# children	# districts	Height	Height z-score
Full sample	13,909	1,728	-0.152	-0.022
			(0.261)	(0.040)
Children with non-white-collar fathers	6,568	1,178	0.377	0.056
			(0.333)	(0.053)
Children with low household expenditure	6,305	1,175	0.048	-0.004
			(0.383)	(0.059)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7. Effects of no school lunch for current and post-junior-high students

Panel (a): Base DID

		BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)
Full sample	*NoSchoolLunch*	0.085	0.028	0.756	0.006	0.007
(23,472 children, 2,265 districts)	x JuniorHigh	(0.089)	(0.034)	(0.461)	(0.009)	(0.010)
	NoSchoolLunch	0.023	0.001	0.489	0.006	0.010
	x PostJH	(0.102)	(0.04)	(0.524)	(0.011)	(0.011)
Children with non-white-collar fathers	*NoSchoolLunch*	0.380***	0.143***	2.002***	0.031**	0.026*
(11,744 children, 1,657 districts)	x JuniorHigh	(0.132)	(0.05)	(0.691)	(0.013)	(0.014)
	NoSchoolLunch	0.262*	0.092	1.524**	0.031**	0.025
	x PostJH	(0.152)	(0.058)	(0.769)	(0.016)	(0.016)
Children with low household	*NoSchoolLunch*	0.349**	0.136**	2.101***	0.041***	0.046***
expenditure	x JuniorHigh	(0.153)	(0.057)	(0.783)	(0.015)	(0.015)
(10,534 children, 1,547 districts)	*NoSchoolLunch*	0.339**	0.139**	2.122**	0.022	0.027
	x PostJH	(0.166)	(0.065)	(0.832)	(0.018)	(0.018)

Panel (b): DID with IPTW and propensity-score trimming

		BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)
Full sample	*NoSchoolLunch*	0.038	0.020	0.573	0.000	0.005
(17,987 children, 1,729 districts)	x JuniorHigh	(0.110)	(0.042)	(0.563)	(0.011)	(0.011)
	NoSchoolLunch	0.141	0.060	1.185*	0.010	0.022*
	x PostJH	(0.131)	(0.051)	(0.672)	(0.013)	(0.013)
Children with non-white-collar fathers	*NoSchoolLunch*	0.513***	0.202***	2.791***	0.031*	0.040**
(8,105 children, 1,121 districts)	x JuniorHigh	(0.177)	(0.067)	(0.908)	(0.018)	(0.019)
	NoSchoolLunch	0.363*	0.142*	2.139**	0.020	0.028
	x PostJH	(0.200)	(0.076)	(1.025)	(0.019)	(0.019)
Children with low household	*NoSchoolLunch*	0.451**	0.189***	2.501***	0.047**	0.052***
expenditure	x JuniorHigh	(0.177)	(0.066)	(0.904)	(0.018)	(0.018)
(7,571 children, 1,119 districts)	*NoSchoolLunch*	0.661***	0.260***	3.467***	0.032	0.041*
	x PostJH	(0.209)	(0.082)	(1.068)	(0.024)	(0.022)

Note: Standard errors clustered at the district level are in parentheses. *PostJH* is a dummy variable for 15- to 17-year-old students who are not junior-high students. Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendices

Appendix 1. Details of the Japanese School Lunch Program

Trends in school lunch provision

Official statistics on trends in school lunch provision are drawn from the School Lunch Data Book (SLDB) for 1978–2004 and the Current Status Survey on School Meals (CSSSM) for 2006–2015. The SLDB is published by the National School Health Center of Japan, and the CSSSM is conducted by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT).[1] We calculate annual nationwide school lunch coverage rates as the fraction of municipal school students with school lunch in all municipal school students.

Municipalities provide either "complete school lunch," "complementary school lunch," or "milk-only school lunch." "Complementary school lunch" provides dishes except for staple food, and students bring steamed rice or bread on their own. "Milk-only school lunch" only provides milk, and students bring their food. We use the sum of "complete school lunch" and "complementary school lunch" to calculate the school lunch coverage rate, though the latter has always been less than one percent for both elementary and junior-high students.

Figure A1 shows trends in the national school lunch coverage rate for municipal junior-high students and in the coverage rate in our sample (as defined in Subsection 4.2). Both lines show similar increasing trends, although the latter is always higher than the former. This difference can be explained by low school lunch coverage rates in prefectures excluded from our sample, as detailed in Appendix 3. The school lunch coverage rate for municipal elementary school students (not shown in the figure) has always been above 98 percent since 1978.

Revisions in nutritional standards

[1] www.mext.go.jp/b_menu/toukei/chousa05/kyuushoku/1267027.htm

Nutritional standards in the Japanese school lunch program have been revised eight times since 1954. During our study period (1975–1994), revision occurred only once in 1986. In the 1986 revision, three-year age categories for elementary school children were replaced by two-year categories, target values for fat (17g, 20g, and 24g for age groups 6–8, 9–11, and 12–14, respectively) were replaced with the maximum percentage of energy intake from fat of 30 percent, and the target amounts were slightly reduced for energy (e.g., from 850kcal to 820kcal for ages 12–14) and protein (e.g., from 36g to 32g for ages 12–14) (Nozue, 2011). The revision did not cause significant changes in the energy or fat contained in school lunch (Narusaka, 1996).

Optional school lunch programs

Since the 1990s, the Ministry of Education has allowed municipalities to make their school lunch programs optional as a temporary measure for schools that have newly started school lunch programs (Asahi Shimbun, 1996). Under an optional lunch program, students have a choice between a school lunch and a home-prepared lunch, which might hinder the estimation of causal effects of having school lunch on child outcomes. During our study period 1975–1994, however, only a handful of municipalities had optional programs, and we remove all children who could have optional school lunch from our sample, as detailed in Appendix 3.

Appendix 2. Details of Inverse Probability of Treatment Weighting (IPTW) and Propensity-Score Trimming

As a robustness check, we use IPTW and propensity-score trimming in the DID analysis to balance municipal characteristics between the treatment and control groups in order to make the common trend assumption more plausible. The propensity scores for IPTW and trimming are estimated using a logistic regression model, in which *NoSchoolLunch* dummy is regressed on district characteristics. The unit of observation is a child. The set of explanatory variables closely resembles those in the district level regression described in Subsection 4.4 but, following Imbens (2015), we use numerical variables instead of dummy variables when possible, omit insignificant variables, and add interaction terms of significant variables to increase the fitness of the model. Specifically, we control prefectural population density and a linear time trend. We also control for the logged median population size for each municipal population size category obtained from Statistics Bureau, Ministry of Internal Affairs and Communications (2012).

IPTW gives each subject a weight equal to the inverse of the probability of treatment the subject received. Thus, for students in district *d* the weight equals $\frac{1}{\mathrm{p}_d}$ if *d* is a treatment district and $\frac{1}{1-\mathrm{p}_d}$ if *d* is a control district, where p_d denotes the propensity score of a lack of school lunch in *d*. Propensity-score trimming excludes observations based on the propensity scores to ensure sufficient overlap in characteristics between the treatment and control districts. We define the common support as the interval between the 1st and 99th percentiles of the estimated propensity scores and exclude from both the treatment and control groups observations that are outside of the common support (Stuart, 2010). We also exclude observations whose estimated propensity score is smaller than 0.1 or larger than 0.9 (Imbens, 2015).

In these procedures, we require the absolute normalized difference of all the explanatory variables used in the DID analysis between the treatment and control districts to be less than 0.25, as values above 0.25 are considered problematic (Rubin, 2001). Table A1 compares normalized differences of individual and district characteristics between the treatment and control groups before and after IPTW and trimming for the full sample and the two low-SES subsamples. Normalized differences of individual characteristics (i.e., control variables in DID analysis) are shown in Panel (a) and those of district characteristics are shown in Panel (b). Both panels indicate that IPTW and trimming reduce the absolute normalized differences substantially. The absolute normalized difference occasionally increases slightly after these procedures, but it is when the original value is small. When the original absolute normalized difference is large, these procedures always reduce it. The absolute normalized differences of individual characteristics are always below 0.25, both with and without IPTW and trimming for all three samples. Some of the raw absolute normalized differences of district characteristics exceed 0.25 without IPTW and trimming, but IPTW and trimming reduce all of them to below 0.25.

Appendix 3. Details of Data Construction and Sample Selection

Identifying elementary and junior-high students

The NNS questionnaire asks if children are in compulsory education but does not distinguish between elementary and junior-high schools. Because a child’s school grade in Japan is strictly determined by the child’s age on April 2nd, we can use children’s age for this purpose and categorize 6- to 11-year-old and 13- to 15-year-old children in compulsory education as elementary and junior-high students, respectively. The school grade of 12-year-olds cannot be determined from the information available in the NNS, but for those who are 12 years old in 1986–1994 we can use the birth month information available in the Comprehensive Survey of Living Conditions (CSLC). The CSLC is an annual large-scale household survey, administered by the MHLW since 1986, and can be merged to our NNS data because the NNS target population is subsampled from that of the CSLC. Because the NNS is conducted in November, we categorize 12-year-olds at the time of survey as elementary students if they are born between April and October, and those born between December and March as junior-high students. We exclude 12-year-olds for 1975–1985 from the sample because we cannot determine their school grade. November-born 12-year-old children in 1986–1994 surveys are also excluded for the same reason.

Sample Exclusion criteria

Our sample consists of 9- to 15-year-old children in elementary and junior-high schools, excluding all the 12-year-olds in the 1975–1985 NNS and the November-born 12-year-olds in the 1986–1994 NNS, as described above. From this sample, we exclude children in prefectures with a high proportion of junior-high students attending non-municipal schools and children in Nagoya city in 1993 and 1994, as detailed below. Next, we limit our sample to children with valid information on height and weight, and exclude a small number of children whose height in *z*-score exceeds 4.0 in absolute value; whose POW is −45 percent or lower (i.e., whose weight is 55 percent or less of the standard weight-for-height); or whose POW is 100 percent or higher (i.e., whose weight is double or more of the standard weight-for-height), as these values may reflect genetic growth disorders or data coding error. These threshold values for POW follow medical guidelines indicating emergency hospitalization for POW of −45 percent or lower (Suzuki, 2016) and metabolic surgery for POW of 100 percent or higher (Kawamura, 1995). A small number of children without a mother in the household or without valid household expenditure data are also excluded. We further limit our sample to children in districts with at least one elementary student and one junior-high student, as required for DID. Lastly, we exclude children in districts with unreliable school lunch information due to conflicting or too few answers, and children in districts where less than half of the elementary school children report having school lunch, as detailed below in this Appendix. Figure A2 presents changes in the sample size after each step of these exclusion criteria.

Reasons for the time frame restriction

The individual-level data from the NNS is available for every year since 1975. We do not use the NNS data collected after 1994 for two reasons. First, after 1994 the food diary in the NNS covers only one day of each household's choice from weekdays and Saturdays. Because our data do not contain information on the day of the week and school lunch is not served on Saturdays, our imputed school lunch status will be understated due to the families who chose Saturday for the food diary. Second, the expansion of optional school lunch programs since the late 1990s obscures the causal interpretation of our estimate, as detailed in Appendix 1 and in the next section.

Excluding municipalities with optional school lunch programs

As described in Appendix 1, municipalities that newly start school lunch programs after the 1990s are allowed to make their school lunch program optional and give students a choice between school lunch and home-prepared lunch (Asahi Shimbun, 1996). This is potentially a concern to our analysis because optional programs may attenuate our treatment-effect estimate. Among municipalities included in our data, the city of Nagoya started an optional program on trial at seven of its municipal junior highs in 1993 (Asahi Shimbun, 1993). We identify children living in Nagoya in 1993 and 1994 using prefecture and municipal population size information and exclude those children from our final sample. Two cities in the Chiba prefecture, Funabashi and Matsudo, also started optional school lunch programs before 1994 (Asahi Shimbun, 1993), but this is not a concern for our sample because our sample does not include children in the Chiba prefecture due to its high proportion of junior-high students attending non-municipal schools, as explained below.

Municipal provision of school lunch

From the individual reports, we only know whether a child had school lunch on certain days, and this information does not necessarily accurately reflect municipal school lunch provision for the following reasons. First, some children might miss school lunch due to sickness or extra-curricular activities such as excursions. Second, children might attend municipal schools outside of the municipality. Third, a non-negligible number of junior-high students attend private and national schools, and few of these schools provide school lunch. We presume that the first type of measurement error is small because we use the food diary of three consecutive days and because the NNS instructs households to choose a survey period that reflects the household's usual diet and avoid a period with a special event. The second type of measurement error is also expected to be small due to the strict Japanese school district system. Attending municipal schools outside of the municipality of residence is only permitted for special reasons such as geographic difficulties in commuting to the designated school (Nakamura, 2000). Some parents make a false resident registration so that their children can attend municipal schools outside of the school district (e.g., Mainichi, 2015), but this is illegal and uncommon. To address the third type of measurement error, we exclude prefectures where the proportion of junior-high students attending non-municipal schools is high. Because information on the school ownership type is not collected in the NNS, we identify prefectures with a high proportion of junior-high students attending non-municipal schools using the number of junior-high students by prefecture and ownership type from the School Basic Survey conducted by MEXT.[2] Considering that the proportion of junior-high students attending non-municipal schools has increased over time, we exclude ten prefectures where the proportion is five percent or higher in 1994: Tokyo, Kochi, Nara, Kanagawa, Kyoto, Hyogo, Hiroshima, Osaka, Chiba, and Mie (listed in the descending order of the proportion). This exclusion drops 35.4 percent of the observations.

[2] http://www.mext.go.jp/b_menu/toukei/chousa01/kihon/1267995.htm

We argue that our imputed school lunch status data are reasonably accurate due to the exclusion of these prefectures, the use of the majority rule, and the exclusion of districts with insufficient information. Additionally, we complement our imputation with official statistics from the School Lunch Data Book (SLDB). Specifically, we regard children in prefectures with 99 percent or higher school lunch coverage, according to the SLDB, as having school lunch. For the years without SLDB statistics, we apply the same imputation using linearly interpolated values.

Additionally, the comparison of the derived school lunch status with the official statistics confirms the reliability of our method. As described in Appendix 1, municipalities' school lunch programs provide either "complete school lunch," "complementary school lunch," or "milk-only school lunch." According to SLDB, 57.6–66.9 percent of Japanese municipal junior-high students had "complete school lunch," 0.5–0.7 percent had "complementary school lunch," and 18.9–26.5 percent had "milk only school lunch" during our study period. Different types of school lunch contents obscure the causal interpretation of "having school lunch." We use "complete school lunch" and "complementary school lunch" as the definition of school lunch in our analysis, but not "milk-only school lunch" because of its limited scope. However, the NNS does not include information on the type of school lunch, and hence whether a respondent's report on school lunch in the NNS complies with our conceptual notion of school lunch warrants further investigation. To examine which type of school lunch is captured in our school lunch status imputed from the NNS, we conduct a prefecture-year level regression analysis in which we regress the imputed school lunch status on the officially reported participation rates of the three types of school lunch. The dependent variable is the prefecture-year proportion of junior-high students who report having school lunch in our sample. In constructing this variable, we relax the last two exclusion criteria in Figure A2 and include districts with too few and/or conflicting reports and districts where less than half of elementary students report having school lunch. From SLDB, we draw data on the prefecture-year proportions of municipal junior-high students who have "complete," "complementary," and "milk-only" school lunch. We then conduct a regression analysis in which we regress the variable constructed from our sample on the three SLDB variables.

The results are shown in Table A2. The coefficients on the proportions of students with “complete” and “complementary” school lunch are both significantly positive and close to one, implying that the vast majority of these students report having school lunch. The coefficient of the proportion of students with milk-only school lunch is not significant and is small in magnitude, implying that few of these students report having school lunch. These findings indicate that the school lunch information in the NNS data closely coincides with the sum of “complete school lunch” and “complementary school lunch,” which is our definition of school lunch.

Appendix 4. Additional Tables

Summary statistics of the district-level data

Table A3 shows the summary statistics of the district-level data. Junior-high lunch is less common in large municipalities. About 28 percent of the districts with junior-high lunch are in cities with a population of 150,000 or more, whereas about 54 percent of the districts without junior-high lunch are in cities with a population of 150,000 or more.

Coefficients on year dummies in NoSchoolLunch Logit regression

Table A4 presents the estimated coefficients on year dummies of the *NoSchoolLunch* regression, which are not reported in Table 4. Consistent with the overtime increase in school lunch coverage (Figure A1), the estimated coefficients are significantly negative for most years since 1983.

Appendix 5. Further Robustness Checks

Permutation test

To address a concern in our DID analysis that correlation among children close in age and location might lead to underestimation of standard errors, we implement non-parametric permutation tests in the spirit of Bertrand et al. (2004) and Abadie et al. (2010). In this test, the treatment *NoSchoolLunch* is randomly assigned to control districts holding the same frequency as in the original sample, and the base DID model is estimated based on the hypothetical treatment assignment. Repeating this procedure many times yields the distribution of the estimated placebo treatment effect. The resulting placebo distributions from 1,000 random draws are shown in Figure A3 for (a) the full sample, (b) the subsample of children with non-white-collar fathers, and (c) the subsample of children with low household expenditure in Panels (a) to (c), respectively. The vertical red lines indicate the actual estimate, i.e., the estimated treatment effect in the original estimation. Numbers below each graph show the actual estimate and the pseudo *P*-value, defined as the frequency ratio of placebo estimates exceeding the actual estimate in absolute terms. The results are highly consistent with our main results from the base DID presented in panel (b) of Table 5. For BMI, BMI *z*-score, POW, and the two obesity measures, the actual estimates in all three samples are always in the right tail of the distribution, and in both of the low-SES subsamples the pseudo *P*-values are all smaller than 0.05, with the exception of one value of 0.068, while the pseudo *P*-values in the full sample are all larger than 0.1. For both of the underweight measures, the actual estimates are around the middle of the distribution, and the pseudo *P*-values are all larger than 0.1 in all three samples. These findings imply that the statistical significance of our estimated school lunch effects is not due to the misspecification of correlation structure.

Adding interaction terms with JuniorHigh dummy

To address the potential violation of the common trend assumption, we test a model with the interaction terms of the *JuniorHigh* dummy with five-year period dummies, municipal size dummies, and prefectural population density. These interaction terms are to account for the possibility that changes in the growth patterns after entering junior highs due to lifestyle alterations, such as time spent on studying and school sports club activities, differ by urbanicity and survey timing. Because the regression analysis of the *NoSchoolLunch dummy* reveals the significant effects of urbanicity and survey years on school lunch status (Table 4), our results might be confounded by the differential effects of attending junior highs on body weight. The estimated effects for the base DID and DID with IPTW and propensity-score trimming are shown in Panels (a) and (b) of Table A5, respectively. The results are highly similar to those from the base models (Table 5), confirming that our results are robust to this modification.

Similarly, we add the triple interactions of the *JuniorHigh* dummy, year dummies, and prefecture dummies to the DID regression model to allow for differential effects of attending junior high by prefecture-year. The estimated effects for the base DID and DID with IPTW and propensity-score trimming are shown in Panels (a) and (b) of Table A6, respectively. Again, this modification causes little change to the results, confirming the robustness of our results.

Synthetic control method

While the common trend assumption implies similar pre-treatment trends between the treatment and control groups, mean BMI by age and gender presented in Figure 4 show small but noticeable differences between groups in pre-treatment ages. To address this concern, we implement a synthetic control method proposed by Abadie et al. (2010) in which we create an artificial control group whose pre-treatment trends are more similar to the treatment group than the original control group. Using aggregate panel data at the level of the treatment group and control subgroups, this method assigns weights to the control subgroups so as to minimize the differences in predictors between the treatment group and the weighted average of the control subgroups prior to the treatment, where predictors could include pre-treatment outcome values and other covariates. Using this weight, an artificial control group called the synthetic control group is constructed as the weighted average of the control subgroups. Based on the premise that the post-treatment path of this synthetic control group approximates the counterfactual path of the treatment in the absence of the treatment, the post-treatment outcome trends are compared between the treatment and synthetic control groups. To obtain statistical inference of treatment effect estimates, Abadie et al. (2010) also propose a placebo test that treats each of the control subgroups as a placebo treatment group and applies the synthetic control method that regards the actual treatment group and the rest of the control subgroups as placebo control subgroups.

We implement the synthetic control method to estimate the effect of the lack of school lunch on age-specific means of BMI, BMI *z*-score, and POW, as well as rates of obesity and underweight under IOTF and POW definitions. Because this method requires control subgroups with a sufficiently large number of observations, we group *NoSchoolLunch* districts by five periods (1975–78, 79–82, 83–85, 86–89, and 90–94) and two regions (east Japan consisting of prefectures Nagano, Niigata, Shizuoka, Yamanashi, and those to the east of these, and west Japan consisting of prefectures to the west of these) to create ten control subgroups. Using this grouping, we construct aggregate panel data at the level of the treatment group and control subgroups over the ages of 7 to 11, 13, and 14. We exclude 12-year-olds and 15-year-olds due to the small number of observations. To ensure a sufficiently long pre-treatment period, we add 7- and 8-year-old elementary students. For predictors, we use pre-treatment values of the outcome variable and maternal BMI; and for obesity and underweight outcomes, we also use pre-treatment values of BMI *z*-score for IOTF-based outcomes and those of POW for POW-based outcomes. Among predictors, the pre-treatment outcome values are averaged for ages 7–10 and age 11, and the other predictors are averaged over the entire pre-treatment ages of 7 to 11.

Figure A4 presents (a) the results for the full sample, (b) the subsample of children with non-white-collar fathers, and (c) the subsample of children with low household expenditure in Panels (a)-(c), respectively. In the graphs on the left, black lines plot the gap in the mean between the actual treatment group and synthetic control group against age, and gray lines show gaps between the placebo treatment group and the corresponding synthetic control group (henceforth, the trend graph). The vertical red line indicates age 12, the threshold age that divides elementary and junior-high students. The graphs on the right present the histogram of the ratio of post- to pre-treatment mean-squared prediction error (MSPE) for the actual and placebo treatment groups, where bars with vertical red dotted line indicate the actual value and other bars indicate the ten placebo values. Loosely speaking, if the actual value is larger than the ten placebo values, the null hypothesis of zero treatment effect can be rejected at the 10 percent level. Overall, the results are consistent with the DID results shown in the main text. In the full sample, the trend graphs show relatively small changes in the actual gap, and the histograms do not support significant post-treatment changes. An exception is the obesity rate under POW, with its histogram indicating a significant post-treatment change and the trend graph showing a moderate increase after age 11. In the subsample of children with non-white-collar fathers, the trend graphs show clear increases in the actual gap after age 11 for BMI, BMI *z*-score, POW, and the obesity rates under both definitions. The histograms also indicate significant post-treatment changes for all of these outcomes with the exception of the BMI *z*-score, where one placebo value exceeded the actual value. For the underweight rate under POW definition, the actual gap for the underweight rate under IOTF definition decreases after age 11, and according to the histogram this post-treatment change is statistically significant, whereas almost no changes in the actual gap are observed in the trend graph and the histogram does not support significant post-treatment changes. In the subsample of children with low household expenditure, post-treatment increases are observed in the trend graphs for BMI,

BMI *z*-score, POW, and the obesity rates under both definitions, and the histograms indicate significant post-treatment changes for all of these outcomes except for POW. For both of the underweight measures, the post-treatment changes in the actual gap appear small in the trend graphs and are not statistically significant according to the histograms.

Relaxing the prefecture selection criterion

Children in prefectures with 5 percent or more of junior-high students attending non-municipal (i.e., private or national) schools as of 1994 are omitted from our main analysis. As a robustness check, we relax this selection criterion to include prefectures where the proportion is between 5 and 10 percent. The majority of the previously excluded prefectures meet this criterion (Chiba, Hiroshima, Hyogo, Kanagawa, Kyoto, Mie, Nara, and Osaka), and the prefectures excluded under this new criterion are only Kochi (15.6 percent) and Tokyo (22.9 percent). The estimated treatment effects from the DID regression are shown in Table A7. The estimated effects are highly similar to those with our preferred selection criterion shown in Table 5.

Exclusion of 12- and 15-year-olds

Among 12-year-olds, some are elementary and others are junior-high students, depending on the birth month. Similarly, among 15-year-olds, some are current and others are post-junior-high students. Hence, our results might be affected by potential differential effects of the birth month on children's height and weight between the treatment and control groups. To address this concern, we estimate the DID models excluding 12- and 15-year-olds. The estimated treatment effects from the DID regression, shown in Table A8, are highly similar to those without this exclusion shown in Table 5.

To assess the effects of the birth month on growth patterns, Figure A5 plots year-adjusted means of height by gender over age in month for the treatment and control groups for the full sample and the two low-SES subsamples, where the age is aggregated by quarter of year. Children surveyed before 1986 and November-born children are excluded because we cannot determine their age in month. To adjust for differences in survey year composition by age, we subtract the gap between the mean by year and the aggregate mean from each value, as we do for Figure 4. The vertical gray dotted lines show the birth months during the second quarter, i.e., April through June. Because children born on April 2nd are the eldest in class and those born on April 1st are the youngest, under the Japanese age-grade system, if parents select the birth timing to exploit the relative age effect, the selection effect would likely lead to an increased difference in height between children born in the first and second quarters. For all three samples and for both genders, the influence of the birth timing appears small, and little systematic difference in the birth month effects is observed between the treatment and control groups.

References

Abadie A, Diamond A, Hainmueller J, 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105: 493–505.

Asahi Shimbun, 1993, October 27. *The Optional School lunch programs at junior high Schools in Nagoya Had a 60% Utilization Rate, Falling below Expectations*. Retrieved from Asahi Shimbun Kikuzo database (In Japanese).

Asahi Shimbun, 1996, August 27. *Shaken by O157: Turning Point for Uniformly Imposed School Lunch*. Retrieved from Asahi Shimbun Kikuzo database (In Japanese).

Bertrand M, Duflo E, Mullainathan S, 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119 (1): 249–275.

Imbens GW, 2015. Matching methods in practice: Three examples. *Journal of Human Resources* 50 (2): 373–419.

Kawamura I, 1995. Surgical treatment of obesity. *Journal of the Japanese Society of Internal Medicine* 84 (8): 1295–1299 (in Japanese).

Mainichi, 2015, December 10. *Himeji City Board of Education to Warn 18 Children Suspected for Attending a School out of School District within This Year.* Retrieved from mainichi.jp (in Japanese).

Nakamura M, 2000. School choice in Japan: Learning from the experiences of other countries. *Nagano Technical College Research Reports* 34: 107–116.

Narusaka M, 1996. Changes in fat supply in school lunch menus in Okayama City. *Japanese Journal of Nutrition and Dietetics* 54 (2): 121–128 (in Japanese).

Nozue M, 2011. The contribution of school lunch to dietary intakes in school children and application of dietary reference intakes: A case study of fifth-grade school children. *Ph.D. Thesis, Kagawa Nutrition University* (in Japanese).

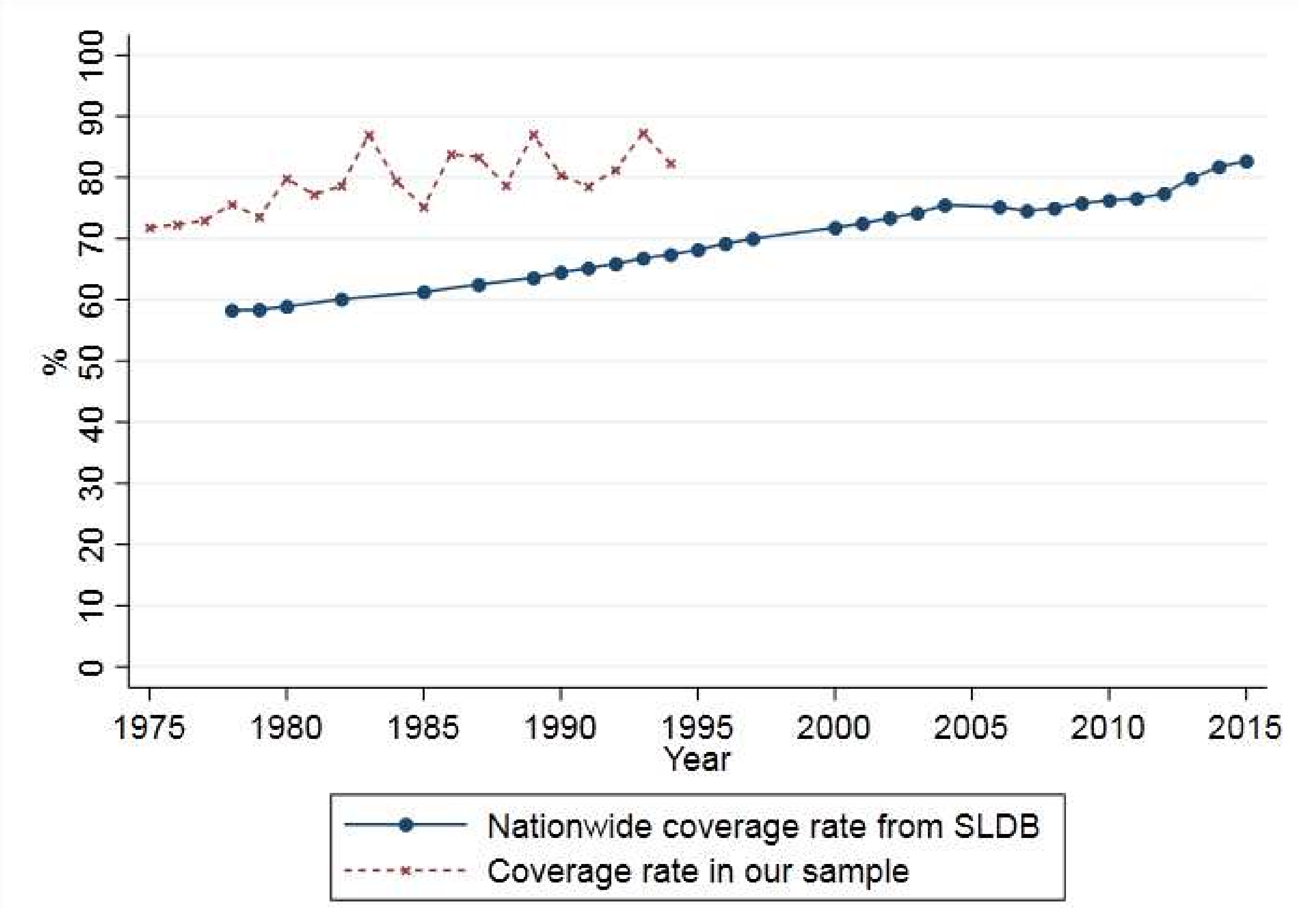
Rubin DB, 2001. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services & Outcomes Research Methodology* 2: 169–188.

Statistics Bureau, Ministry of Internal Affairs and Communications (2012) Historical statistics of Japan. https://www.stat.go.jp/naruhodo/dl/zuhyou/c1s_origin/y680203000.xls

Stuart EA, 2010. Matching methods for causal inference: A review and a look forward. *Statistical Science* 25 (1): 1–21.

Suzuki, MH, 2016. Endocrine and metabolic emergencies: Points of initial management. *Journal of the Japanese Society of Internal Medicine* 105 (4): 676–682 (in Japanese).

Figure A1. Trends in the school lunch coverage rate for municipal junior-high students



Note: The nationwide coverage rate is based on the School Lunch Data Book (SLDB) and defined as the share of municipal junior-high students who have school lunch. The sample coverage rate is the share of junior-high students in districts with school lunch in our sample. For the discussion on the difference between the two coverage rates, see Appendix 1.

Figure A2. Sample exclusion criteria and changes in the sample size

Initial sample: 9- to 15-year-old children in compulsory education, excluding 12-year-olds in the 1975–1985 NNS and November-born 12-year-olds in the 1986–1994 NNS

5,580 districts, 37,990 children

↓ Exclude children in prefectures with a high proportion of junior-high students in non-municipal schools and children in the city of Nagoya in 1993 and 1994

3,280 districts, 24,530 children

↓ Exclude children without valid information on height and weight

3,228 districts, 22,241 children

↓ Exclude children with deviant height and weight

3,227 districts, 22,205 children

↓ Exclude children without a mother in the household or without valid household expenditure data

3,205 districts, 21,327 children

↓ Exclude districts without elementary students and districts without junior-high students

2,661 districts, 20,064children

↓ Exclude districts with conflicting or too few answers on school lunch

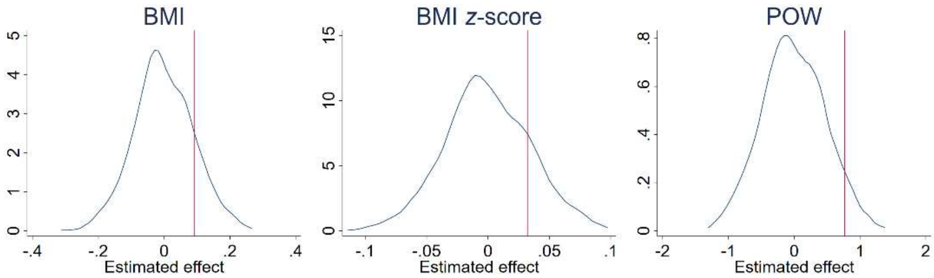
2,311 districts, 18,669 children

↓ Exclude districts without school lunch for elementary school children

2,265 districts, 18,271children

Figure A3. Permutation tests: empirical distributions of placebo estimates

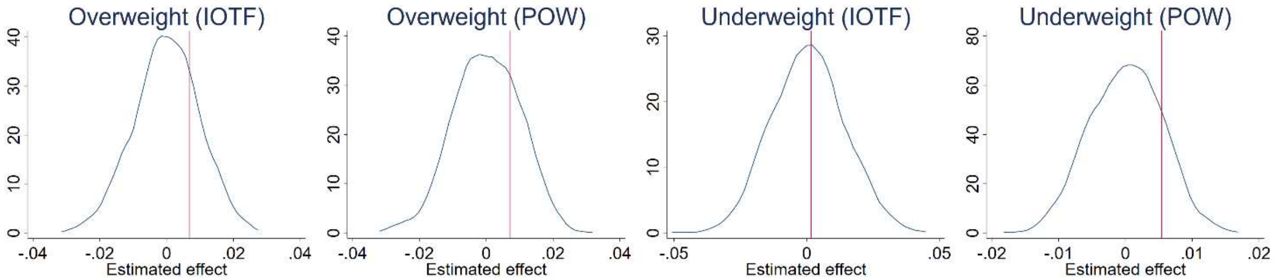
Panel (a): Full sample



actual estimate=0.093, pseudo *P*-value=0.296

actual estimate=0.032, pseudo *P*-value=0.351

actual estimate=0.767, pseudo *P*-value=0.109



actual estimate=0.007, pseudo *P*-value=0.464

actual estimate=0.007, pseudo *P*-value=0.5

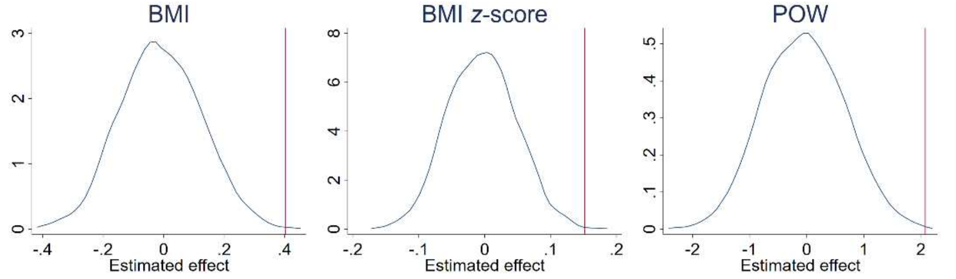
actual estimate=0.002, pseudo *P*-value=0.901

actual estimate=0.005, pseudo *P*-value=0.334

Note: Each graph shows the kernel density estimate of the distribution of estimated placebo treatment effects using Epanechnikov kernel. The vertical red line represents the actual estimate. "Pseudo *P*-value" refers to the frequency ratio of placebo estimates exceeding the actual estimate in absolute terms (in 1,000 random draws).

Figure A3 (cont.)

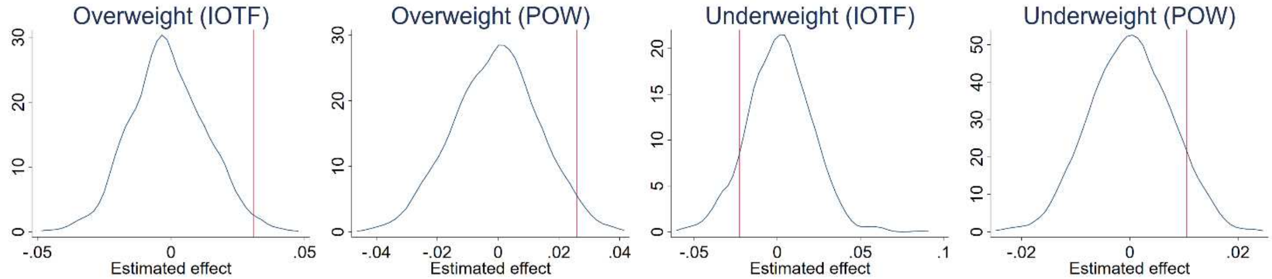
Panel (b): Children with non-white-collar fathers



actual estimate=0.403, pseudo *P*-value=0.001

actual estimate=0.152, pseudo *P*-value=0.002

actual estimate=2.076, pseudo *P*-value=0.002



actual estimate=0.031, pseudo *P*-value=0.032

actual estimate=0.026, pseudo *P*-value=0.068

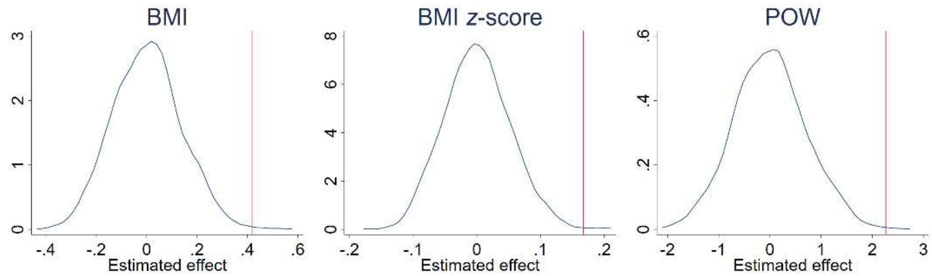
actual estimate=-0.023, pseudo *P*-value=0.22

actual estimate=0.01, pseudo *P*-value=0.164

Note: Each graph shows the kernel density estimate of the distribution of estimated placebo treatment effects using Epanechnikov kernel. The vertical red line represents the actual estimate. "Pseudo *P*-value" refers to the frequency ratio of placebo estimates exceeding the actual estimate in absolute terms (in 1,000 random draws).

Figure A3 (cont.)

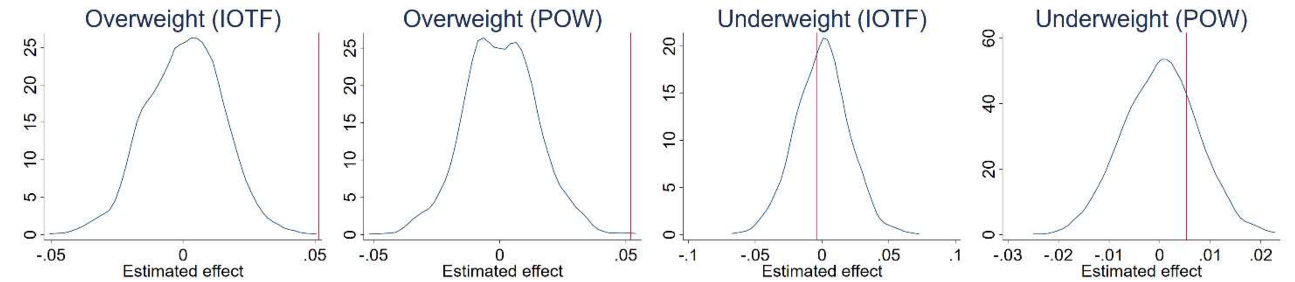
Panel (c): Children with low household expenditure



actual estimate=0.418, pseudo *P*-value=0.004

actual estimate=0.167, pseudo *P*-value=0.003

actual estimate=2.266, pseudo *P*-value=0.001



actual estimate=0.051, pseudo *P*-value=0

actual estimate=0.052, pseudo *P*-value=0

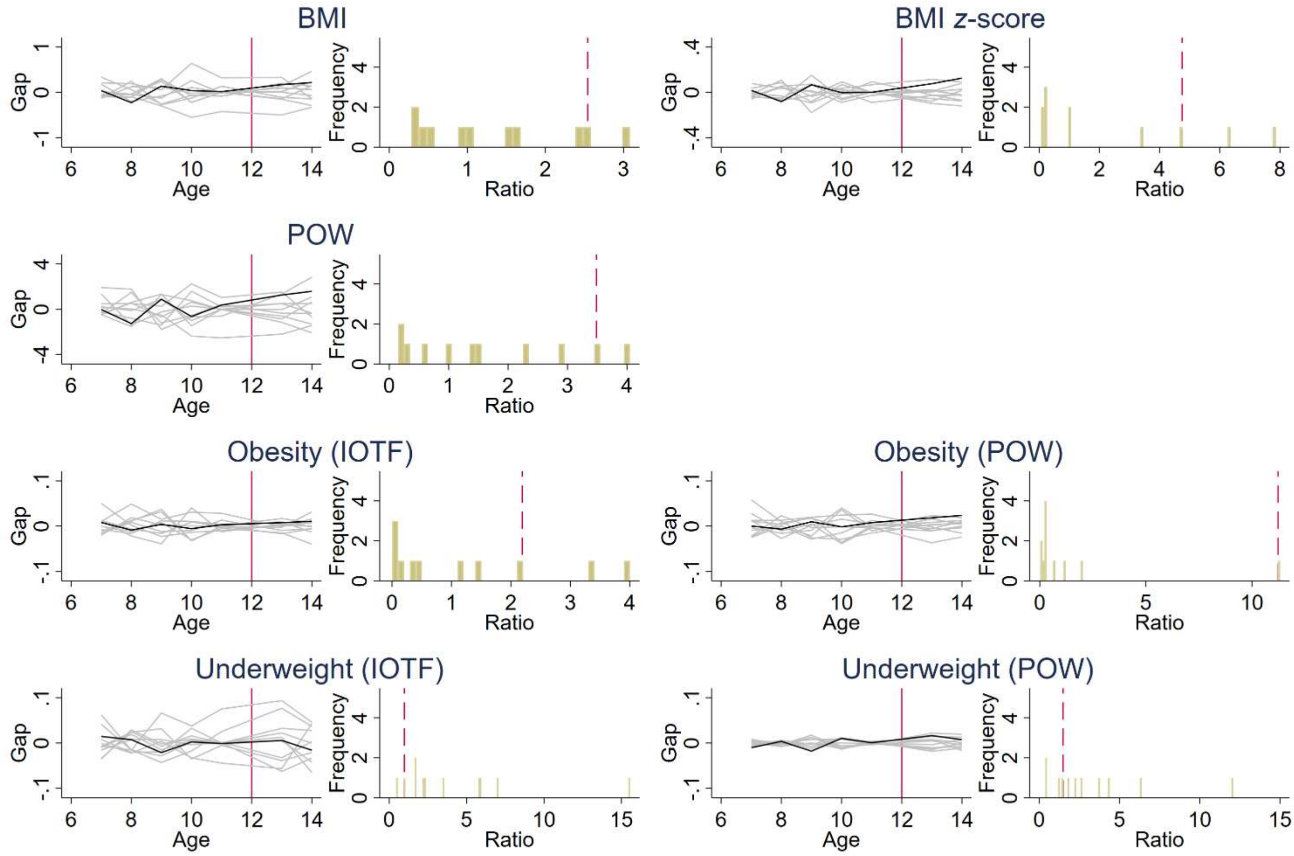
actual estimate=-0.004, pseudo *P*-value=0.821

actual estimate=0.005, pseudo *P*-value=0.48

Note: Each graph shows the kernel density estimate of the distribution of estimated placebo treatment effects using Epanechnikov kernel. The vertical red line represents the actual estimate. “Pseudo *P*-value” refers to the frequency ratio of placebo estimates exceeding the actual estimate in absolute terms (in 1,000 random draws).

Figure A4. Results from the synthetic control analysis

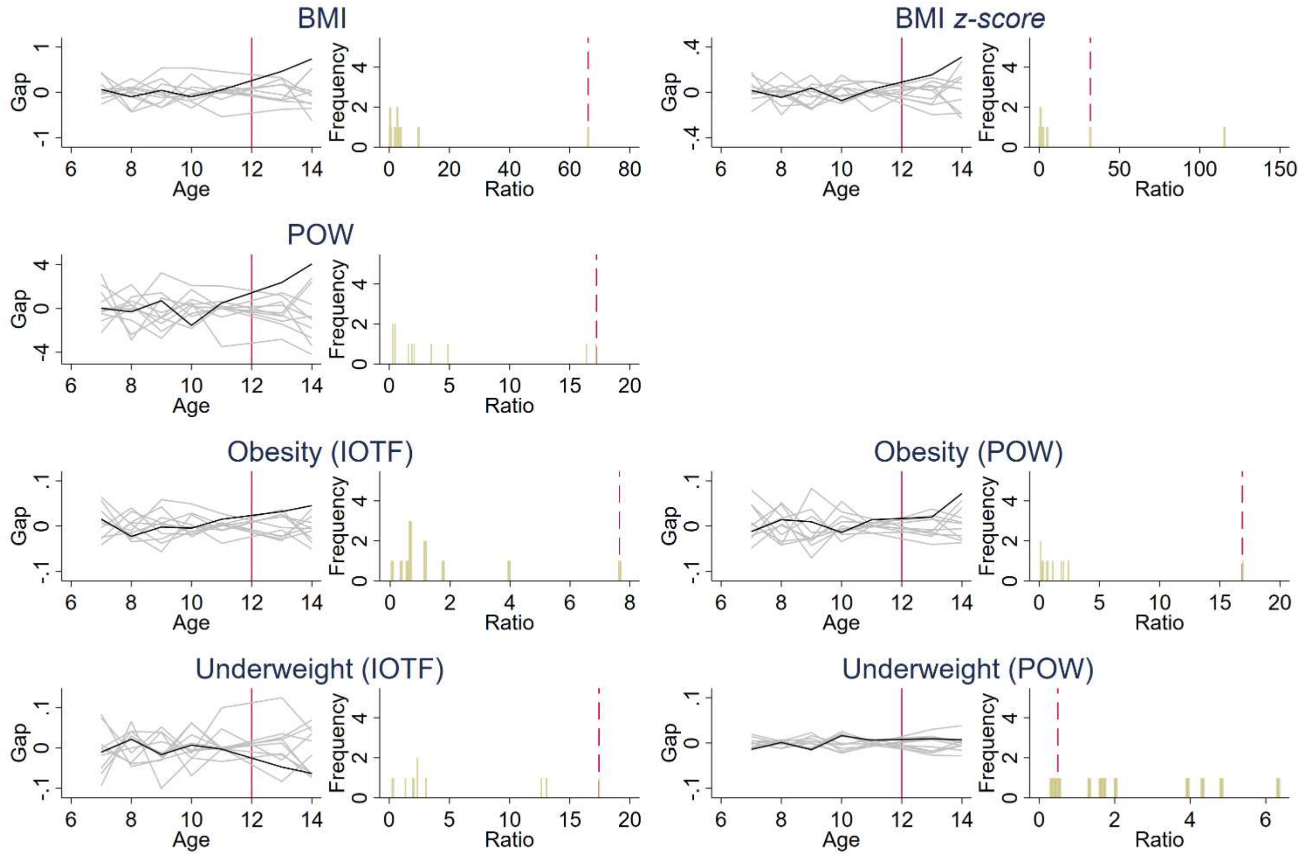
Panel (a): Full sample



Note: In the graphs on the left, black lines plot the gap in the mean between the actual treatment group and synthetic control group against age, and gray lines show gaps between the placebo treatment group and the corresponding synthetic control group. Vertical red lines indicate age 12, the threshold age that divides elementary and junior-high students. The graphs on the right present the histogram of the ratio of post- to pre-treatment MSPE (mean squared prediction error) for the actual and placebo treatment groups, where bars with a vertical red dotted line indicate the actual value and other bars indicate the ten placebo values.

Figure A4. (cont.)

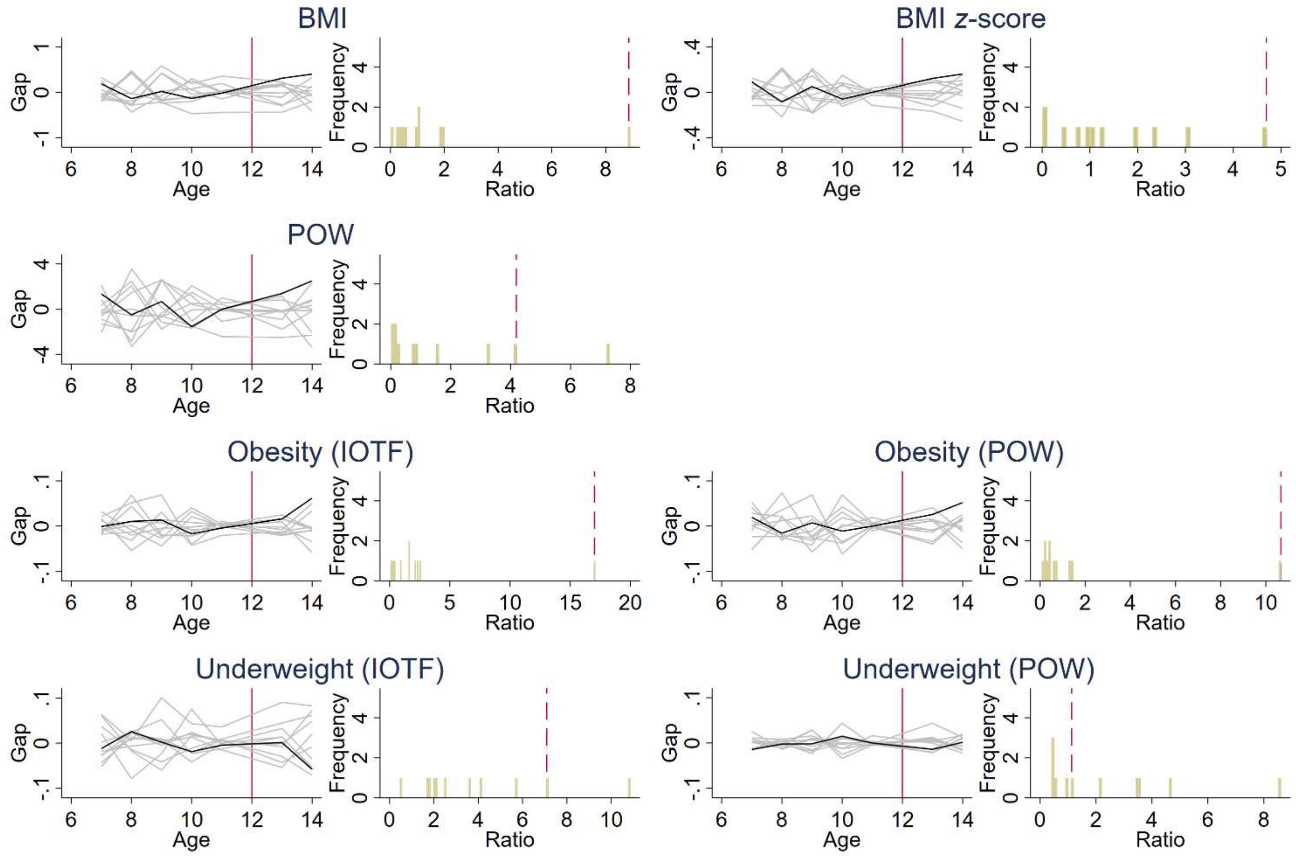
Panel (b): Children with non-white-collar fathers



Note: In the graphs on the left, black lines plot the gap in the mean between the actual treatment group and synthetic control group against age, and gray lines show gaps between the placebo treatment group and the corresponding synthetic control group. Vertical red lines indicate age 12, the threshold age that divides elementary and junior-high students. The graphs on the right present the histogram of the ratio of post- to pre-treatment MSPE for the actual and placebo treatment groups, where bars with a vertical red dotted line indicate the actual value and other bars indicate the ten placebo values.

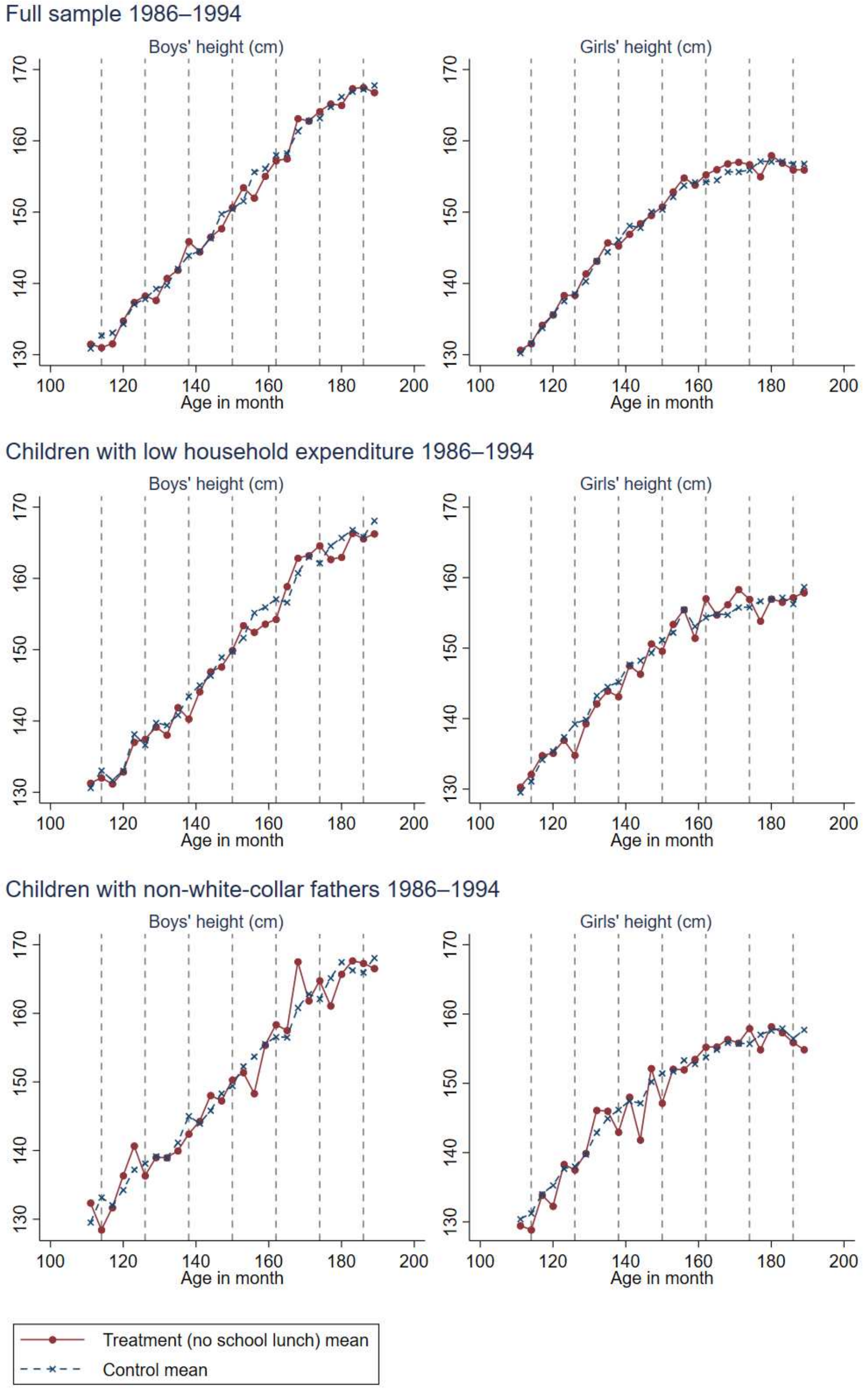
Figure A4. (cont.)

Panel (c): Children with low household expenditure



Note: In the graphs on the left, black lines plot the gap in the mean between the actual treatment group and synthetic control group against age, and gray lines show gaps between the placebo treatment group and the corresponding synthetic control group. Vertical red lines indicate age 12, the threshold age that divides elementary and junior-high students. The graphs on the right present the histogram of the ratio of post- to pre-treatment MSPE for the actual and placebo treatment groups, where bars with a vertical red dotted line indicate the actual value and other bars indicate the ten placebo values.

Figure A5. Mean height over age in month



Note: The age is aggregated by quarter of year. Children surveyed before 1986 or born in November are excluded for data reasons. To adjust for differences in survey year composition by age, we subtract the gap between mean by year and aggregate mean from each value. The vertical gray dotted lines indicate the birth months during the second quarter.

Table A1. Normalized differences of individual and district characteristics

Panel (a): Individual characteristics (control variables in DID analysis)

Variable	Full sample		Children with non-white-collar fathers		Children with low household expenditure	
Trimming and IPTW-weighting	No	Yes	No	Yes	No	Yes
Male	0.014	0.012	0.028	0.053	0.013	-0.004
Age	-0.025	-0.024	-0.045	-0.043	0.003	0.013
Father's age	0.004	-0.026	0.000	-0.052	0.004	0.010
Father's height (*z* score by age, sex, and 5-year cohort)	0.072	0.030	0.041	-0.003	0.048	0.036
Father's BMI (*z* score by age, sex, and 5-year cohort)	-0.035	-0.006	-0.031	-0.035	-0.028	-0.022
Father's height and BMI missing	0.009	-0.002	-0.033	-0.069	0.017	-0.023
Father: white-collar worker (the reference category)	0.027	-0.011			0.013	-0.027
Father: laborer	0.011	0.004	0.060	0.018	0.025	-0.001
Father: self-employed	0.016	-0.019	0.028	-0.059	0.023	0.000
Father: agriculture/fisheries/forestry	-0.104	0.022	-0.131	0.063	-0.096	0.043
Father: other occupation (not working)	0.002	-0.021	-0.002	-0.020	0.009	-0.011
Without father in household	0.005	0.026			0.007	-0.001
Mother's age	-0.001	0.011	-0.043	-0.048	-0.026	-0.009
Mother's height (*z* score by age, sex, and 5-year cohort)	0.056	-0.008	0.003	-0.038	0.028	-0.011
Mother's BMI (*z* score by age, sex, and 5-year cohort)	-0.097	-0.059	-0.073	-0.055	-0.097	-0.059
Mother's height and BMI missing	-0.001	0.021	-0.007	0.029	-0.013	0.004
Mother: white-collar worker (the reference category)	0.024	0.072	0.059	0.072	0.041	0.069
Mother: laborer	-0.020	0.045	-0.015	0.058	0.010	0.045
Mother: self-employed	0.019	0.004	0.019	-0.001	0.039	0.055
Mother: agriculture/fisheries/forestry	-0.150	-0.042	-0.167	-0.021	-0.168	-0.053
Mother: other occupation (not working)	0.067	-0.074	0.076	-0.088	0.048	-0.101
Grandfather in household	-0.101	0.021	-0.123	0.001	-0.073	0.059
Grandmother in household	-0.099	0.035	-0.080	0.030	-0.050	0.093
# of children in household (below 18 years old)	-0.042	0.002	-0.029	-0.028	-0.007	0.017
Per-member household expenditure (defined as 100% = families with lowest expenditures in each survey year)	-0.147	0.038	-0.172	-0.035	-0.186	-0.093

Note: None of the absolute normalized differences exceed 0.25. Occupational proportions refer to the proportion of 23- to 54-year-old workers in each occupation.

Table A1. (cont.)

Panel (b): District characteristics

Variable	Full sample		Children with non-white-collar fathers		Children with low household expenditure	
Trimming and IPTW-weighting	No	Yes	No	Yes	No	Yes
Year	-0.222	0.061	-0.246	0.003	-0.105	0.095
Prefectural population density (1,000 person/km2)	0.063	-0.126	0.082	-0.071	0.056	-0.151
Logged municipal population size	0.729*	-0.048	0.794*	-0.021	0.686*	-0.095
District: mean child height (*z* score)	0.111	0.124	0.102	0.065	0.124	0.091
District: mean child BMI (*z* score)	-0.031	0.044	-0.037	-0.002	0.009	0.051
District: child obesity rate (IOTF)	-0.004	0.064	0.018	0.046	-0.023	0.001
District: child underweight rate (IOTF)	0.034	-0.028	0.054	-0.026	-0.020	-0.043
District: # of participants	-0.145	-0.075	-0.106	-0.093	-0.112	-0.039
District: mean age	-0.080	0.170	-0.100	0.116	-0.044	0.238
District: median per-member household expenditure	-0.328*	-0.041	-0.346*	-0.097	-0.353*	-0.093
District: mean household size	-0.212	-0.011	-0.207	-0.025	-0.195	0.020
District: proportion of white-collar worker	0.059	-0.051	0.008	-0.131	0.094	-0.065
District: proportion of laborer	0.021	0.039	0.070	0.062	0.044	0.019
District: proportion of self-employed	0.124	0.024	0.165	0.028	0.110	0.021
District: proportion of agriculture	-0.214	0.000	-0.213	0.030	-0.215	0.028
District: proportion of working women	-0.143	0.190	-0.133	0.197	-0.115	0.239

Note: * indicates absolute normalized difference exceeding 0.25. Occupational proportions refer to the proportion of 23- to 54-year-old workers in each occupation.

Table A2. Prefecture-year level regression of the proportion of junior-high students who had school lunch

Variable	Coefficients	
SLDB proportion of students with complete school lunch	0.893***	(0.092)
SLDB proportion of students with complementary school lunch	0.947**	(0.458)
SLDB proportion of students with milk only	0.031	(0.107)
Constant	8.083	(8.863)
N	427	
Adjusted R^2	0.320	

Note: Standard errors are in parentheses. The unit of observation is a prefecture-year. The proportion of junior-high students who had school lunch in our NNS sample is calculated at the prefecture-year level and regressed on the proportions of those with the three types of school lunch in the official SLDB statistics. Both the regressand and the regressors are in percentage terms. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A3. Summary statistics of the district-level data

Variable	Districts with no junior-high lunch Mean	Std. Dev	Control districts Mean	Std. Dev	
District: mean child height (*z* score)	0.015	0.360	-0.010	0.366	
District: mean child BMI (*z* score)	0.001	0.395	0.010	0.370	
District: child obesity rate (IOTF BMI 25+)	0.162	0.142	0.159	0.134	
District: child underweight rate (IOTF BMI 18.5-)	0.118	0.115	0.113	0.108	
District: # of participants	81.27	27.49	85.4	26.87	***
District: proportion of age 1–19	0.316	0.067	0.309	0.070	*
District: proportion of age 20–39 (reference)	0.263	0.075	0.260	0.076	
District: proportion of age 40–59	0.277	0.071	0.275	0.071	
District: proportion of age 60+	0.144	0.079	0.156	0.089	***
District: median per-member household expenditure	0.503	0.184	0.571	0.189	***
District: mean household size	4.152	0.582	4.291	0.630	***
District: proportion of white-collar worker (reference)	0.386	0.207	0.370	0.208	
District: proportion of laborer	0.346	0.192	0.345	0.188	
District: proportion of self-employed	0.209	0.152	0.187	0.153	***
District: proportion of agriculture/fisheries/forestry	0.059	0.141	0.098	0.173	***
District: proportion of working women	0.573	0.188	0.608	0.197	***
Prefectural population density (1,000 person/km^2)	0.510	0.433	0.462	0.446	**
Municipal size: 11 largest cities	0.183	0.387	0.031	0.173	***
Municipal size: cities with 150k+ population	0.362	0.481	0.248	0.432	***
Municipal size: cities with 50–150k population	0.224	0.417	0.217	0.412	
Municipal size: cities with 50k- population	0.096	0.295	0.103	0.303	
Municipal size: towns & villages (reference)	0.134	0.341	0.402	0.490	***
Region block: Hokkaido & Tohoku	0.171	0.377	0.184	0.387	
Region block: Kanto	0.081	0.273	0.190	0.392	***
Region block: Chubu (reference)	0.595	0.491	0.485	0.500	***
Region block: Kinki	0.064	0.245	0.018	0.132	***
Region block: Chugoku & Shikoku	0.090	0.286	0.123	0.329	**
Region block: Kyushu & Okinawa	0.252	0.434	0.194	0.395	***
Year	1982.7	5.6	1984.2	5.5	***
Number of districts	469		1,648		

Note: Occupational proportions refer to the proportion of 23- to 54-year-olds in each occupation. The number of districts is smaller than that in the full sample DID regression because 148 districts with less than five children of age 1 to 11 are removed from this sample. *, **, and *** indicate statistically significant differences in means between the control and treatment districts at the 10%, 5%, and 1% levels, respectively.

Table A4. The estimated year effects from Logit regression of *NoSchoolLunch*

Variable	Model1 (N=2,117)		Model2 (N=2,075)	
Year 1976	0.056	(0.299)	0.089	(0.335)
Year 1977	0.089	(0.319)	0.200	(0.377)
Year 1978	-0.248	(0.318)	-0.138	(0.366)
Year 1979	-0.252	(0.334)	-0.179	(0.364)
Year 1980	-0.424	(0.319)	-0.359	(0.364)
Year 1981	-0.318	(0.336)	-0.335	(0.385)
Year 1982	-0.536	(0.333)	-0.344	(0.374)
Year 1983	-1.135***	(0.384)	-1.127**	(0.439)
Year 1984	-0.849**	(0.346)	-0.746*	(0.392)
Year 1985	-0.710**	(0.345)	-0.642	(0.401)
Year 1986	-1.129***	(0.340)	-1.057***	(0.391)
Year 1987	-1.002***	(0.357)	-0.755*	(0.409)
Year 1988	-0.788**	(0.338)	-0.684*	(0.398)
Year 1989	-1.395***	(0.414)	-1.365***	(0.466)
Year 1990	-0.958***	(0.368)	-0.970**	(0.440)
Year 1991	-0.718**	(0.356)	-0.415	(0.408)
Year 1992	-0.844**	(0.371)	-0.729*	(0.431)
Year 1993	-1.312***	(0.418)	-1.153**	(0.466)
Year 1994	-1.203***	(0.448)	-1.048**	(0.522)
Region block dummies	Yes		No	
Prefecture dummies	No		Yes	

Note: This table presents the year effects estimated in the *NoSchoolLunch* Logit regression, which are not reported in Table 4. The reference year is 1975. Standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A5. Effects of no school lunch: Interacting *JuniorHigh* with urbanicity variables and period dummies

Panel (a): DID + interaction terms

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	18,271	2,265	0.048	0.020	0.526	0.003	0.008	0.007	0.010*
			(0.094)	(0.036)	(0.491)	(0.010)	(0.010)	(0.015)	(0.006)
Children with non-white-collar fathers	9,226	1,657	0.402***	0.156***	2.111***	0.035**	0.032**	-0.017	0.016*
			(0.145)	(0.056)	(0.760)	(0.015)	(0.015)	(0.021)	(0.009)
Children with low household expenditure	8,317	1,542	0.390**	0.158***	2.023**	0.053***	0.054***	0.004	0.011
			(0.155)	(0.059)	(0.795)	(0.015)	(0.015)	(0.023)	(0.008)

Panel (b): DID with IPTW and propensity-score trimming + interaction terms

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	13,909	1,728	0.108	0.043	0.873	0.004	-0.002	0.002	0.003
			(0.105)	(0.040)	(0.540)	(0.011)	(0.011)	(0.015)	(0.006)
Children with non-white-collar fathers	6,568	1,178	0.546***	0.211***	2.873***	0.039**	0.039**	-0.029	0.010
			(0.167)	(0.064)	(0.869)	(0.017)	(0.017)	(0.023)	(0.009)
Children with low household expenditure	6,305	1,175	0.507***	0.206***	2.645***	0.057***	0.053***	-0.009	0.007
			(0.168)	(0.064)	(0.851)	(0.016)	(0.017)	(0.023)	(0.008)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A6. Effects of no school lunch: Interacting *JuniorHigh* with state-year fixed effects

Panel (a): DID + interaction terms

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	18,271	2,265	0.074	0.031	0.611	0.004	-0.002	0.018	0.003
			(0.106)	(0.041)	(0.555)	(0.011)	(0.011)	(0.016)	(0.006)
Children with non-white-collar	9,226	1,657	0.438***	0.182***	1.943**	0.029*	0.017	-0.027	-0.002
fathers			(0.168)	(0.065)	(0.881)	(0.016)	(0.017)	(0.026)	(0.010)
Children with low household	8,317	1,542	0.502***	0.217***	2.883***	0.055***	0.044**	-0.006	0.005
expenditure			(0.192)	(0.072)	(0.986)	(0.018)	(0.019)	(0.027)	(0.010)

Panel (b): DID with IPTW and propensity-score trimming + interaction terms

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	13,909	1,728	0.093	0.033	0.719	0.003	0.000	0.034*	0.001
			(0.128)	(0.049)	(0.674)	(0.014)	(0.014)	(0.019)	(0.007)
Children with non-white-collar	6,568	1,178	0.645***	0.258***	2.641**	0.054***	0.036*	-0.013	0.002
fathers			(0.204)	(0.080)	(1.086)	(0.020)	(0.021)	(0.031)	(0.013)
Children with low household	6,305	1,175	0.670***	0.294***	3.766***	0.071***	0.047**	0.007	0.001
expenditure			(0.236)	(0.088)	(1.214)	(0.022)	(0.021)	(0.033)	(0.011)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A7. Effects of no school lunch: Relaxing the prefecture selection criterion

Panel (a): DID

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	24,460	3,148	0.011	0.005	0.307	0.003	0.006	0.001	0.006
			(0.067)	(0.026)	(0.351)	(0.007)	(0.007)	(0.011)	(0.004)
Children with non-white-collar fathers	11,723	2,179	0.255**	0.092**	1.426***	0.030***	0.028***	-0.009	0.012*
			(0.105)	(0.041)	(0.552)	(0.011)	(0.011)	(0.016)	(0.007)
Children with low household expenditure	11,076	2,086	0.274**	0.118***	1.557***	0.036***	0.043***	-0.015	0.003
			(0.111)	(0.042)	(0.577)	(0.011)	(0.011)	(0.017)	(0.006)

Panel (b): DID with IPTW and propensity-score trimming

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Children with non-white-collar fathers	9,274	1,686	0.323**	0.128**	1.955***	0.032**	0.038***	-0.002	0.015*
			(0.130)	(0.050)	(0.682)	(0.014)	(0.014)	(0.019)	(0.008)
Children with low household expenditure	9,695	1,770	0.339**	0.146***	2.046***	0.043***	0.056***	-0.004	0.006
			(0.139)	(0.053)	(0.726)	(0.014)	(0.014)	(0.021)	(0.007)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. In Panel (b), the results for the full sample are omitted because IPTW and propensity-score trimming failed to reduce the normalized differences of individual and district characteristics between the treatment and control groups below the threshold value of 0.25. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A8. Effects of no school lunch: Excluding 12-year-olds and 15-year-olds

Panel (a): DID

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	15,245	2,168	0.152	0.051	1.065**	0.013	0.013	-0.008	0.007
			(0.101)	(0.039)	(0.524)	(0.010)	(0.011)	(0.015)	(0.006)
Children with non-white-collar fathers	7,580	1,504	0.601***	0.220***	3.104***	0.042***	0.036**	-0.043**	0.008
			(0.154)	(0.059)	(0.805)	(0.015)	(0.016)	(0.022)	(0.009)
Children with low household expenditure	6,718	1,380	0.504***	0.199***	2.696***	0.057***	0.056***	-0.026	0.003
			(0.172)	(0.066)	(0.886)	(0.017)	(0.017)	(0.023)	(0.008)

Panel (b): DID with IPTW and propensity-score trimming

Sample	# children	# districts	BMI	BMI *z*-score	POW	Obesity (IOTF)	Obesity (POW)	Underweight (IOTF)	Underweight (POW)
Full sample	11,831	1,665	0.183	0.073	1.236**	0.008	0.012	-0.019	0.003
			(0.120)	(0.046)	(0.613)	(0.013)	(0.013)	(0.018)	(0.005)
Children with non-white-collar fathers	5,471	1,069	0.683***	0.256***	3.530***	0.045**	0.046**	-0.038	0.008
			(0.191)	(0.073)	(0.971)	(0.019)	(0.020)	(0.026)	(0.009)
Children with low household expenditure	5,422	1,104	0.705***	0.284***	3.691***	0.068***	0.068***	-0.032	0.010
			(0.189)	(0.072)	(0.953)	(0.018)	(0.019)	(0.026)	(0.008)

Note: Standard errors clustered at the district level are in parentheses. For the list of control variables included in each regression, see Table 3. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.