Policy Diffusion Through Election

By

Hitoshi Shigeoka
(Simon Fraser University, The University of Tokyo, IZA and NBER)

Yasutora Watanabe
(The University of Tokyo)

August 23, 2023

CREPE DISCUSSION PAPER NO. 142

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

Policy Diffusion Through Elections*

Hitoshi Shigeoka†
Simon Fraser University,
University of Tokyo,
IZA and NBER

Yasutora Watanabe‡
The University of Tokyo

June 30, 2023

Abstract

Staggered difference-in-differences designs are pervasive in policy evaluations but little is known about the mechanisms of policy diffusion: How and why do such policies spread across jurisdictions? In this study, we highlight the role of elections in policy diffusion in settings where municipal elections are asynchronous due to historical reasons. First, we empirically show the presence of policy diffusion using neighbors' election cycles as instruments for neighbors' policy adoption. Second, we further demonstrate interactions of municipalities' election cycles with neighbors' adoption and show that they follow neighbors' policy only during their own election timing, indicating that policy diffuses through elections.

Keywords: policy diffusion, elections, subsidy, child healthcare, political budget cycles
JEL codes: D04, D78, H73, H75

---

* The authors would like to thank Edward Glaeser, Emir Kamenica, Hiroyuki Kasahara, Keisuke Kondo, Hikaru Ogawa, Katumi Shimotsu, Raphael Parchet, and seminar participants at AMES 2022, EMES 2022, GRIPS, Hitotsubashi University, JSQPS 2023 Winter Meeting, Keio University, Kyoto Summer Workshop on Applied Economics, Kyoto University, LACEA-LAMES 2022, NBER Japan Meeting 2022, Nihon University, Rochester University, SAET 2021, for their valuable suggestions. Shigeoka gratefully acknowledges the financial support from Abe fellowship. Watanabe acknowledges the financial support from JSPS KAKENHI Grant Number 21H05062. Yusuke Ishihata, Eiken Kawa, Kento Kitajima, Ryoki Maekawa, Hiromi Matsukiyo, Genta Okada, Naoya Sho, Hideaki Takahashi, Masahiro Tamae, Kotaro Tsukada, Keisuke Ueda, Futo Watanuki, and Masahiro Yamada provided excellent research assistance. All remaining errors are solely our responsibility.

† Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A1S6, CANADA, The University of Tokyo, IZA, and NBER. Email: hitoshi_shigeoka@sfu.ca

‡ Graduate School of Economics and Graduate School of Public Policy, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, JAPAN. Email: yasutora@e.u-tokyo.ac.jp

1. Introduction

Economists have evaluated the impact of policies by exploiting the differential timing of policy adoption across jurisdictions (staggered difference-in-differences designs). However, little is known about the underlying processes and mechanisms of policy *diffusion*—how and why such policy adoptions spread across jurisdictions in particular ways. The mechanism, especially the one affecting the timing of policy adoption, may influence the research design and interpretation of impact evaluations. However, examining the determinants of policy diffusion is challenging because neighbors' policy adoption can be endogenous (Manski 1993). For example, neighboring municipalities may suffer from a common policy problem, and thus, decide to adopt a similar policy at a similar time. Alternatively, common interest groups may pressure adjacent municipalities to implement similar policies (Gibbons and Overman 2012).

To overcome these endogeneity issues, we adopt a novel approach by utilizing the institutional features of municipal elections in Japan, where the *timing* of municipal elections is exogenously different across municipalities. Immediately after WWII, all municipal elections were held on the same day in April 1947. Given the four-year term of the mayors, subsequent elections were scheduled every fourth year in April. However, by the start of our dataset in April 2005, a large fraction of municipal elections was no longer held at the time of these nationwide synchronized elections for idiosyncratic historical reasons, such as the deaths of mayors that occurred in the past 60 years.[1]

This unique setting allows us to use neighbors' election cycles as an instrument for neighbors' policy adoption. The exclusion restriction of this instrument is likely to be satisfied because a mayor's death multiple decades ago, for example, would not have substantially influenced citizens' and politicians' behavior in the 2000s. Indeed, we empirically demonstrate that election timing between two adjacent municipalities is orthogonal. The relevance of our instruments follows the idea of political budget cycles (PBC, hereafter), which document that policy adoption is concentrated at the timing of the election (Nordhaus 1975; Rogoff and Sibert 1988; Rogoff 1990).[2]

---

[1] Our setup sharply contrasts with the cases of simultaneous/synchronized elections (e.g., US state elections). However, unsynchronized elections are not uncommon. For example, Indian state elections (Khemani 2004; Cole 2009), German local elections across states (Foremny and Riedel 2014; Englmaier and Stowasser 2017), and Italian municipal elections (Repetto 2018) are not synchronized.
[2] See Drazen (2001), and de Haan and Klomp (2013) for reviews of the literature on PBC.

In terms of policy, we focus on municipal subsidies for child healthcare—where generosity is measured by the maximum eligibility age for the subsidy—as a key example of the provision of public services at the local level. In the last decade, Japanese municipalities have rapidly expanded subsidies for child healthcare, and there have been substantial variations in generosity across municipalities and over time. This specific spending level is suitable for studying policy diffusion for several reasons.

First, the policies are highly visible to voters. Municipalities typically provide coverage until the ages of 6, 12, and 15 years, usually a multiple of three, corresponding to the start of primary school, secondary school, and high school in Japan. These discrete numbers are easily comprehensible for voters. Second, policies are easily comparable for both politicians and electorates. They can easily recognize that own jurisdiction (i)'s adopted policy is inferior to that of their neighbor (j). For instance, with discrete numbers, it is evident that the coverage in municipality i with a subsidy for up to six years old looks relatively less generous than that of the neighboring municipality j with a subsidy for up to nine years old. These features make child healthcare subsidies a populist policy that politicians and voters are concerned about in Japan.[3]

Third, high-frequency policy data at the *monthly* level—which we manually collected for the first time— are available. Such high-frequency data on policy adoption turn out to be vital, as we also find that politicians increase the eligibility age immediately *after* elections, unlike conventional studies in the PBC literature. This effect is masked by the low-frequency data used in the literature, where spending is usually observed at the fiscal year level.[4]

We have three main results. Overall, we provide strong evidence that elections play a significant role in policy diffusion.

First, we demonstrate that municipalities adopt policies—raise the eligibility age around the timing of the elections—which we refer to as the "election timing effect" in the context of child healthcare subsidies in Japan. Compared to the middle two years in 4-year election cycles, a municipality is more likely to increase the eligibility age one year *before* the election,

[3] We refrain from discussing the effectiveness of the policy as it is beyond the scope of this study. See Iizuka and Shigeoka (2018, 2022a), which examined the effect of the same subsidy on healthcare utilization and health outcomes, and showed that the subsidy-induced utilization of healthcare is mostly wasteful.

[4] To the best of our knowledge, the only study that has used the monthly data in PBC literature is Akhmedov and Zhuravskaya (2005), which examined a Russian case. Interestingly, they found a very short-lived increase in government spending just before the election and a decrease right after the election, which also highlights the virtue of high-frequency data. Although they did not consider the neighbors' behavior.

consistent with the PBC. Interestingly, the municipality also expands the coverage one year *after* the previous election. Such a just-after-election pattern disappears for municipalities following uncontested elections, implying that the existence of elections forces politicians to implement the policy immediately after the election, at which point they promise to expand it during the campaign.

Importantly, this result is equivalent to showing that we have a strong first stage in our instrumental variable strategy (i.e., instrument relevance). This election timing effect for own municipality i should also apply to neighbors j; thus, by construction, we use the election cycle of *neighbor* (j) as the IV for the *neighbor*'s policy adoption. Furthermore, because the timing of elections between two adjacent municipalities (i, j) is indeed exogenously different and orthogonal, as shown later, this instrument will likely satisfy the exclusion restriction.

Second, we find solid evidence of policy diffusion using neighbors' election cycles as the instrument for neighbors' policy change. The municipality expands the eligibility age when it is strictly below that of a neighbor, indicating that municipalities try to catch-up and fill the gap with neighbors. The IV estimates for policy diffusion are smaller than the OLS estimates, indicating the endogeneity of neighbors' policy choices due to positively correlated preferences or environments.

Third and most importantly, we show that such policy diffusion concentrates on one's own election timing. When we interact neighbor j's policy adoption with its own i's election cycle, we find that the municipality tries to catch-up with its neighbors *only* at its own election timing. In contrast, we do not observe such behavior during the middle two years. In other words, policy diffusion across municipalities is absent without the influence of elections, suggesting that policy diffusion manifests only *through* elections.

Furthermore, such diffusion is more pronounced for municipalities led by experienced politicians. We find that municipalities led by second-term or more experienced politicians adopt neighbors' policies when their coverage falls behind that of neighbors only around the critical period of their own election cycle. Contrarily, we do not find such a strong diffusion pattern among first-term novice politicians, possibly due to their limited experience in learning the "optimal" timing of policy adoption to catch-up with neighbors.

This study contributes to the literature on policy diffusion (Walker 1969; Mallinson 2021). The literature in economics is mostly limited to tax competition across the US states (Case et al.

1993; Besley and Case 1995; de Paula et al. 2020). DellaVigna and Kim (2022) is a notable exception in studying policy diffusion at the US state level. While descriptive in nature, they document that the best predictor of policy adoption has changed drastically from geographic proximity to political alignment over the last seven decades. This study contributes to the literature by (i) providing a *causal* relationship using a novel identification strategy and (ii) highlighting the importance of elections (cycles) as a determinant of not only policy adoption (of municipality i) but also policy diffusion across jurisdictions (from municipality j to i).

This study also speaks to the underlying mechanisms of policy diffusion. DellaVigna and Kim (2022) summarize that policy diffusion reflects either 1) *common preferences or environments*, 2) *learning* across jurisdictions (Banerjee 1992; Volden et al. 2008), or 3) *competition* among jurisdictions (Besley and Case 1995). Correlated preferences can be ruled out using this novel instrument. While we cannot exclude the possibility of learning, competition across municipalities is likely to be the main channel explaining policy diffusion because politicians' responses are rather strategic and sophisticated: (i) the municipality expands the eligibility age only when its policy is *strictly below* that of a neighbor, but not when it is above or the same; (ii) municipalities try to catch up with neighbors *only* near one's own election; and (iii) such behavior is only observed for *more experienced* politicians in the second-term or higher.

The remainder of this paper is organized as follows. Section 2 describes the institutional background of subsidies and election cycles in Japan, and its datasets. Section 3 provides graphical evidence, and Section 4 presents our identification strategy. Section 5 presents the results, and Section 6 concludes the study.

2. Data and institutional background

This study requires a dataset on the timing of both municipal elections and the municipal adoption of subsidy expansion in Japan. However, a comprehensive database combining information on Japanese municipal elections does not exist. Similarly, a database that combines municipal subsidy information at the *monthly* level in a systematic way does not exist. Therefore, we constructed these datasets from scratch for both explanatory (election cycles) and outcome (subsidy) variables. We manually-collected both datasets from various sources, including municipality web pages, municipal ordinances, local newspapers, historical archives, and other resources in Japan. After collecting the data, we contacted each municipality directly to verify the accuracy of our information.

Consequently, the dataset includes information on both, election and subsidy, for the six most populous prefectures (Saitama, Chiba, Tokyo, Kanagawa, Aichi, and Osaka), covering approximately 300 municipalities.[5] According to national statistics, these six prefectures cover as many as 44.9% of children aged 0–15 years in Japan. We eventually dropped Tokyo (57 municipalities) because its 23 special wards did not hold simultaneous elections in 1947.[6] Consequently, our working sample includes 243 municipalities over ten years between April 2005 and March 2015. The observational unit is each municipality at the monthly level, totaling 120 months. We explain each dataset along with its institutional background in detail below.

2.1. Municipal elections

Japanese local governments consist of prefectures and municipalities.[7] The municipality is the lowest level of jurisdiction. The mayors of the municipalities are elected through simple plurality-rule elections. There are no explicit or implicit term limits for mayors. Most mayors are non-partisan and not subject to the influence of upper jurisdictions (i.e., prefectures). Thus, party affiliation is likely to play little role in this setting, unlike in the US, where party control is an important factor.

In our analysis, we exploit the exogenous variation in election timing to identify the impact of the neighbor's policy.[8] After WWII, all municipal elections were held on the same day in April 1947.[9] Given the four-year term of mayors, subsequent elections were scheduled every fourth year (i.e., 1951, 1955, . . ., 2007, 2011) in April, which is called simultaneous local elections (hereafter "SLE"), if there were no incidents, such as resignation, death, merger, and recall, during the four-year term.

However, by the start of our dataset (i.e., April 2005), a large fraction of municipal elections were no longer held at the time of these nationwide SLEs. Once an election is held off

[5] This includes some municipalities that experience mergers. The results are very similar when we limit our sample to the balanced panels, as shown later.
[6] Our results are qualitatively similar if we add back municipalities in Tokyo to the sample (results available upon request).
[7] There were 47 prefectures (equivalent to states in the US) and 1,719 municipalities (equivalent to counties in the US) in Japan as of January 2015.
[8] This paragraph heavily relies on Fukumoto and Horiuchi (2011) and Fukumoto and Ueki (2015).
[9] Exceptions are the five largest cities, Osaka, Nagoya, Kyoto, Yokohama, and Kobe, which had ayoral elections two weeks before the election of all other municipalities in 1947.

the SLE cycle, subsequent elections remain off the SLE cycle because the subsequent term is always four years and not the remainder of the previous term. For example, in the case of the 2007 SLEs, among the 247 municipalities that we studied, only 21.4% held elections on April 27, 2007. According to Fukumoto and Ueki (2015), municipalities dropped out of SLE cycles because of municipal mergers (42.5%), mostly in the 1950s, followed by resignations (34.0%), deaths (18.2%), and others (5.3%).

Panel A of Figure 1 shows the timing of municipal elections during our sample period of 10 years from April 2005 to March 2015, with a total of 632 elections. Again, while approximately 20% of the municipalities follow SLEs, the rest hold their own elections at different times. As the figure shows, the timing of elections outside SLEs spreads across the years, supporting our claim that deviations from SLEs are idiosyncratic. Indeed, it is difficult to imagine that the factors affecting deviations from SLEs, such as the resignation of mayors and municipal mergers five decades ago, would still substantially influence citizens' and candidates' behaviors in the 2000s.[10]

2.2. Subsidy for child healthcare

We briefly provide a background of the Japanese healthcare system related to this study. The Japanese healthcare system is heavily regulated by the government. Under universal coverage, all citizens must enroll in either an employment-based or residential-based insurance system (Ikegami and Campbell 1995; Kondo and Shigeoka 2013). Regardless of the insurer, individuals face the same fee schedule and benefit packages, both of which are set by the national government.

At the national level, patient cost-sharing—for which the beneficiary is responsible for out-of-pocket—has been set at 30% for people under the age of 70, including all children. Many municipalities provide subsidies for children to cover the remaining costs, which aim to ensure access to essential medical care for children. Children eligible for the subsidy received an

---

[10] To confirm this observation, we conducted a balance test of municipal characteristics for the 2007 and 2011 SLEs held during the sample period. Appendix Table A1 shows that municipal characteristics across the two groups (with and without SLEs) are very similar in both the 2007 and 2011 SLE, and none of the variables included in our regressions are statistically different at the conventional levels. Similarly, Fukumoto and Horiuchi (2016) examined the case of SLE 2003 and conducted a balance test of municipality characteristics between the municipalities that held elections in 2003 SLEs and those that did not hold elections. They found that 14 (7.3%) out of 192 estimates were statistically significant at the conventional five percent levels.

additional insurance card; by simply presenting it, they received a subsidy from medical institutions. Crucially, there is no fiscal externality or benefit spillover; the subsidy from municipality i is *only* available for residents in municipality i. In other words, the children of residents in municipality j who received treatment in hospitals in municipality i did not benefit from the subsidy from municipality i.

To this end, we developed a novel dataset by manually-collecting data on the timing and contents of subsidy expansion at the exact *month* level for ten years (April 2005–March 2015). This dataset is identical to that used in Iizuka and Shigeoka (2018, 2022a, b). Panel B of Figure 1 shows the number of municipalities by the exact adoption timing of subsidy expansion during the sample period, with a total of 606 subsidy expansions. While we see more subsidy expansion in some specific year-month, the timing of adoption is widely spread across the sample period.[11] Figure 2 shows the distribution of municipalities according to the number of subsidy expansions during the sample period. This ranges from zero to seven, with an average of 2.45 expansions per municipality. Only two of the 247 municipalities did not adopt subsidy expansion during our sample period, reassuring us that subsidy expansion is a popular policy and is widespread across almost all municipalities.

Notably, the generosity of the subsidy is mainly reflected by the maximum age until the subsidy is provided (hereafter, the "eligibility age").[12] Figure 3 shows the share of municipalities by eligibility age for outpatient care during the sample period. Note that while the eligibility age is often expressed by school grade (e.g., until the end of junior high school), we loosely use ages throughout this study for convenience, as school grades are almost completely equivalent to age in Japan owing to the strict enforcement of the school entry rule as well as very rare grade retention and advancement rates (Shigeoka 2015). Ages 6, 12, 15, and 18 in Figure 3 correspond to the entry into elementary school, graduation from elementary school, graduation from junior

[11] The small jump in April 2008 is explained by the fact that the central government expanded the eligibility age for the national-level subsidy (i.e., 20% coinsurance rate) from three to six years old (the start of primary school). This national-level subsidy expansion eased the budgetary burden on municipalities, as part of the cost to provide free care for those below six years old was covered by the central government, allowing municipalities to expand coverage to older ages.

[12] There are three other dimensions in subsidy (level of copayment/coinsurance, a refund or in-kind payment, and existence of household income restrictions for subsidy eligibility) but the variations along these dimensions are relatively small (Iizuka and Shigeoka 2022a). Furthermore, politicians exclusively discuss the eligibility age in the official gazette, as shown below.

high school, and graduation from high school, respectively.[13]

Figure 3 shows that subsidies have expanded rapidly to older age groups in the last decade. For example, none of the municipalities adopted a policy with coverage up to the age of 15 in April 2005, the beginning of the sample period. However, this number reached nearly 80% by the end of our sample period, a decade later, in March 2015.[14]

A few more important features of the adoption data should be noted. First, most municipalities stop expansion at age 15, at least in our sample period, which corresponds to the end of junior high school. These ceiling effects should be properly controlled for in the later estimation, as the room for expansion is limited after reaching 15 years, even though municipalities can technically expand their eligibility to higher ages. Second, policy change is always monotonic, as no single municipality lowers the eligibility age in our sample.

This specific spending is suitable for studying policy diffusion. First, the subsidy for child healthcare is a populist policy that both voters and politicians care about, as shown later in gazette, in Figure 5. From the voters' perspective, discrete numbers (e.g., 6, 12, and 15 years) are highly visible and easily comprehensive. From the politician's perspective, it is one of the few policies that mayors can have the discretion to change, as it may only account for approximately 1–2% of the total annual budget of municipalities, which contrasts with policies that target the elderly, which are too costly. Second, a comparison with other municipalities is clear with a discrete number. For example, the coverage in municipality i, with an eligibility age of 6 years, clearly falls behind that of the neighboring municipality j, with an eligibility age of 9 years. Third, high-frequency monthly data are available. To the best of our knowledge, the only study in the PBC literature that uses monthly data is Akhmedov and Zhuravskaya (2005). Note that because data on other spending categories at the monthly level are unavailable, we do not investigate the potential spending spillover to offset the elevated spending on child healthcare subsidies. Fourth, unlike the binary measures of policy adoption studied by DellaVigna and Kim (2022), eligibility age allows us to examine continuous outcome variables with ample variation within and across municipalities.

[13] See Appendix Table A2 for the distribution of eligibility age among 247 municipalities in our sample period.
[14] This figure differs from that of Iizuka and Shigeoka (2018, 2022a) because we dropped Tokyo here, and we did not weigh the number of insurance claims as in Iizuka and Shigeoka (2018, 2022a).

2.3. Descriptive statistics

We construct the final dataset by merging the two datasets on election and subsidy information by municipality and year-month. Then, for each municipality, we merge the information on the bordering neighbors, including their subsidy information and election cycles (our IVs), allowing for multiple observations of neighboring municipalities (j) per municipality (i). The summary statistics of the final dataset are presented in Table 1. Regarding election characteristics, 98% of incumbents were male. The number of terms range from one to ten with an average of approximately two terms as there is no term limit for mayors in Japan. The proportion of mayors in their first term was 39%, and 19% of previous elections were uncontested. In our dataset, 88% of the elections followed the scheduled timing (i.e., a four-year schedule without deviation).

3. Graphical presentation

Before presenting our econometric specifications and results in Sections 4 and 5, we present graphical evidence of the election timing effect of policy adoption in Section 3.1 and then policy diffusion in Section 3.2. Finally, we examine their *interaction* in Sections 4 and 5.

3.1. Election timing effect

Constructing visual evidence for the election timing effect is straightforward. By combining the timing of elections and subsidy expansion from the two figures (Panels A and B of Figure 1), Figure 4 plots the number of subsidy expansions by the time until the next election, measured in months. The vertical line separates the four-year election cycle into distinct year. The far-left interval corresponds to four years before the election (or just after the previous election), and the far-right interval corresponds to the one year before the next election, and there are two middle years in between.

The figure shows two noticeable patterns. First, there are more subsidy expansions one year before the next election than in the middle years, which is consistent with the PBC literature. Second, rather surprisingly, we observe many expansions immediately after the elections, which are similar in magnitude or even larger than such an effect before the election.

We have anecdotal and supportive evidence for such political behavior. Some municipalities mandate that candidates create gazettes summarizing their policies during municipal elections. Many incumbents boast of what they have done in the past to signal their competence. The expansion of subsidy for child healthcare is often included as their accomplishment, like "I have expanded subsidy from age 9 to 12 during my term." However, it is noteworthy that many candidates also list the policies that they claim will be implemented once elected (electoral promises). By definition, opponents can only make promises as they are not in office and thus cannot describe what they have done in the past. However, the incumbent also often posts to-do lists after being elected on the gazette (called "manifesto").

Figure 5 shows an example of this phenomenon. This figure is the official gazette for the municipal election in Tsushima City, Aichi Prefecture, which was held on April 15, 2018. The sentences in the red box indicate subsidy expansion for child healthcare. The candidate on the right is the incumbent (ひび 一昭 in Japanese), who promises to raise the eligibility age for free child healthcare till the end of junior high school (中学卒業), which is equivalent to age 15. The candidate on the left is the opponent (杉山 良介), who also promises exactly the same level (中3) of subsidy expansion. The incumbent won this election and implemented the pledged policy one year later, on April 1, 2019.

At a glance, it may look odd as even though politicians promise, there is no reason to follow the pledged promise and implement it within a year of the election. This finding implies that voters monitor their performance immediately after elections. Interestingly, Panel A of Figure A2 shows that such a pattern immediately after the previous election disappears for politicians who experienced uncontested previous elections, implying that the election itself forces politicians to promise subsidy expansion and eventually adopt the policy immediately after the election.[15]

While this can be particular to the Japanese setting, we show here that low-frequency data on policy adoption cannot detect such political behavior because the annual data cannot distinguish the events that occurred immediately before and after the election. As discussed previously, our advantage is that we have the monthly data on eligibility age. Appendix Figure A1 shows the number of subsidy expansions by year (not month) until the next election as if we had only yearly information about when subsidy expansion was adopted. The figure shows the usual PBC patterns only in the election year as we cannot clearly separate the policies adopted in

[15] Panel B of Figure A2 shows that patterns around the elections look similar between the politicians in the 1st term and 2nd-term or above.

election years into pre- and post-elections.[16]

3.2. Policy diffusion

Next, graphical evidence of policy diffusion across municipalities is provided. Figure 6 shows the case of Saitama Prefecture, located just across the north of Tokyo. This figure demonstrates how subsidies for child healthcare propagate geographically across municipalities via geographical proximity. Each graph describes the adoption status in April from 2005 to 2014. Darker colors indicate that the municipalities have adopted the policy of coverage up to age 15 in the year. The lighter color indicates that the municipalities had already adopted a policy of coverage up to age 15.

The figures show that policy adoption spreads through adjacent municipalities; a municipality is more likely to adopt a policy if nearby municipalities have already implemented it. For example, in 2009, all expansions (darker colors) occurred next to municipalities that had already expanded their eligibility age in the past (lighter color). The year 2010 demonstrates an even stronger pattern of spatial spillovers as subsidy expansion seems to cluster locally. By 2014, the eligibility age in all municipalities in Saitama Prefecture reached 15. This illustration indicates the presence of policy diffusion along a geographical proximity line.

4. Identification strategy

4.1. Empirical challenges

There are three challenges to estimating policy diffusion. First, when is the policy of neighbor j most relevant to the adoption decision in municipality i? In other words, how long does it take for the mayor of municipality i to respond to policy adoption in the neighboring municipality j if he/she wants to respond? Is it three months, six months, or even longer? In Japan, municipal assemblies are held four times per year; thus, the average interval between assemblies is three months. Thus, we start with a lag of three months, assuming that at least three months are necessary to respond to the policy adopted by neighbors. We later experiment by

[16] As exact election dates are often available, some studies try, albeit not perfectly, to distinguish the election held at first half or second half of the year. Specifically, if an election happens in the first half of the year, the election year is regarded as pre-election. Conversely, when the election happens in the second half, the election year is treated as it is (Brender and Drazen 2005).

changing this time lag; however, the results are robust to the choice of the relevant time period.

Second, which neighbor has the most significant influence on municipality i among all neighbors (j)? Based on Besley and Case (1995), voters judge politicians' competence relative to that of their neighboring municipalities. This theory implies that a politician is most influenced by the behavior of those municipalities that their voters judge to be the most salient (Baicker 2005).

We assume that all the bordering neighbors can potentially influence, motivated by the visual observation in Figure 5, but the weight placed on each neighbor (*neighborliness*) can differ. We examine four metrics: "out-migration," "in-migration," "size of the population," and "per capita income." "Out-migration" and "in-migration" determine the degree of neighborliness by the fraction of people that move into (out-migration) or that move from (in-migration) each neighboring municipality. "Size of the population" and "per capita income" computes weight on the difference in population and per capita income between own and neighboring municipalities, reflecting that neighbors with similar population size or per capita income receive more weight.[17] We use the "out-migration" as our baseline. As it is commonly cited that subsidy expansion for child healthcare aims to attract younger parents and boost tax base, it is plausible that mayors are more concerned about the strategies employed by neighboring municipalities to attract residents from their own jurisdiction.[18]

Finally, the biggest challenge is that the neighbors' policy adoption can be endogenous (Gibbons and Overman 2012). For example, neighboring municipalities suffer from common policy problems, such as low fertility rates, and thus decide to expand the subsidy for child healthcare simultaneously. Alternatively, common interest groups may simultaneously pressure nearby municipalities to implement similar policies. If we do not account for such shared

[17] Note that we compute weight so that the weight of bordering neighbors sums up to one.We construct weight (w_{ij}) as follows. For "out-migration" and "in-migration," inter-municipality mobility data are obtained from the 2015 Census. The weight is the fraction of movers from municipality i to j (out-migration) and the fraction of movers from municipality j to i (in-migration). For "Size of the population," the data are obtained from "Sichoson no Sugata," published by the Statistics Bureau (https://www.e-stat.go.jp/regional-statistics/ssdsview, last accessed on August 1, 2019). Following Case et al. (1993) and Baicker (2005), the weight is based on the difference in population size between i and j, or $w_{ij} = 1/\{|\mathrm{Pop}_i - \mathrm{Pop}_j|S_i\}$ where $S_i = \sum_j |\mathrm{Pop}_i - \mathrm{Pop}_j|$. Similarly, for "per capita income," the data are obtained from the same source. The weight is based on the difference in per capita income between i and j or, $w_{ij} = 1/\{|\mathrm{Inc}_i - \mathrm{Inc}_j|S_i\}$ where $S_i = \sum_j |\mathrm{Inc}_i - \mathrm{Inc}_j|$.
[18] Iizuka and Shigeoka (2022a)—using monthly residence information from insurance claim data—show that children (and hence parents) do not move to municipalities with subsidy, suggesting that there are many other reasons (such as school quality) that are more likely to affect the migration decisions.

preferences or environments, we will likely overestimate neighbors' influence owing to a positive correlation in unobserved neighbor characteristics.

To account for this potential endogeneity, we adopt a novel approach by exploiting the fact that the timing of elections across municipalities differs exogenously for idiosyncratic historical reasons. Figure 7 plots the *difference* in the timing of elections between municipality i and adjacent municipality j(s), measured in months. It takes values from 0 to 24 because the difference between the two elections cannot exceed 24 months, given that the election cycle is 48 months.[19] Panel A shows that it is uniformly distributed, indicating that the timing of elections between two adjacent municipalities is indeed exogenously different. In fact, the correlation between the distances in months to the next election between municipalities i and j was as low as 0.0075. Panel B, which excludes SLEs, displays a similar pattern. Hence, we use neighbors' election cycles as an instrument for their policy levels.

4.2. Estimating equation

For municipality i, whose neighboring municipality is j(s), the main specification is written as

$$A_{it} = \sum_{k\neq 2,3}^{k=1,4} \alpha_{-k} E_{it}^{-k} + \beta 1\left(A_{i\tilde{t}} < A_{j\tilde{t}}\right) + \sum_{k\neq 2,3}^{k=1,4} \rho_{-k}\left\{E_{it}^{-k} \times 1\left(A_{i\tilde{t}} < A_{j\tilde{t}}\right)\right\} + \gamma A_{it-1} + \delta X_{it}' + \theta_i + \pi_{tp} + \varepsilon_{it}, \ [1]$$[20]

where A_{it} is the eligibility age for the subsidy at time t (in months), and $A_{i\tilde{t}}$ and $A_{j\tilde{t}}$ are analogously defined for i and j at $\tilde{t} = t - 3$ (i.e., a three-month lag as previously discussed). E_{it}^{-k} $(k = 1, 4)$ is a dummy that takes the value one if the year is k years before the next election. As election cycles are usually fixed every four years, we can, in principle, treat them as exogenous. To be conservative, we construct E_{it}^{-k} based on the number of years until the next *expected* election, following Khemani (2004) and Cole (2009). The results are almost identical if we use the years until the next *actual* election because nearly 90% of elections follow as

[19] If the difference between the two election cycles of municipality i and j is 33 months, this can be viewed as a difference of 15 months, given that the election cycle is 48 months.
[20] The alternative model is the hazard specification, but we do not adopt such an approach because applying instrumental variables in the non-linear hazard model is not straightforward, and additionally our outcome of maximum eligibility age can occupy various values (see Appendix Table A2); thus, the setting does not fit well to discrete-choice decision to adopt a policy in hazard model.

scheduled during our sample period (See Table 1). The reference years are the two middle years between the elections.[21] $1(A_{i\tilde{t}} < A_{j\tilde{t}})$ is a dummy that takes the value one if the eligibility age in municipality i is strictly below that of municipality j. The discreetness of the eligibility age allows us to define this variable without measurement errors.

We are particularly interested in whether municipalities care *more* about neighbors' policy levels during their own election timings. Thus, we also include the interaction terms between the election cycle dummies and a dummy that takes the value one if the eligibility age is below that of the neighbors ($E_{it}^{-k} \times 1(A_{i\tilde{t}} < A_{j\tilde{t}})$).

α_{-k} ($k = 1, 4$) captures the effect of election cycles relative to the two middle years in the absence of policy diffusion. β captures the policy diffusion in the absence of the effect of election timing effects. Our main coefficient of interest is ρ_{-k} ($k = 1, 4$), which captures the interaction effect of election timing and policy diffusion.

We include the lagged eligibility age (A_{it-1}) to capture the monotonicity and ceiling effects of subsidy expansion, as described in Section 2.2. Particularly, as most municipalities stop expanding subsidies at age 15, the room for expansion is substantially different at ages 6 and 12. A_{it-1} captures these heterogeneous effects. However, the inclusion of a lagged variable (A_{it-1}) introduces known mechanical endogeneity issues. As our panel is relatively long, we estimate Equation [1] using a standard fixed effects estimator. Using Arellano-Bond (1991) type GMM estimators yields similar results (results available upon request).

We include municipality FE (θ_i), which captures any time-invariant municipality characteristics, such as preferences for more generous policies for children (e.g., childcare). We also include the year-month FE, which is allowed to differ by prefecture p (π_{tp}). Such flexible fixed effects capture other policies or economic shocks common across all municipalities within the same prefecture to mitigate concerns about shared preferences or the environment. The vector X_{it}' includes both mayor- and municipality-level controls. Mayor-level controls include the gender and terms of the incumbents.[22] Municipality-level controls include the fraction of the population aged 0-15, 15-64, population density, and log income per capita, whereas all

[21] Moreover, we separate the middle two years into each year, but the coefficient of our interests just before and after the elections are quantitively very similar.

[22] To construct these mayor-level controls, we additionally collected information on one election before our sample started in April 2005.

municipality-level controls are available only at the yearly level. ε_{it} is the error term. To account for serial correlation within the municipalities, standard errors are clustered at both municipality i and neighboring municipality j levels.[23]

As $1(A_{i\tilde{t}} < A_{j\tilde{t}})$ is potentially endogenous to the outcome of our interest, we instrument it by the timing of the neighbor j's and own i's election cycle dummies, $E_{j\tilde{t}}^{-k}$ and $E_{i\tilde{t}}^{-k}$ ($k = 1, 4$), as well as lagged eligibility ages, $A_{j\tilde{t}-1}$ and $A_{i\tilde{t}-1}$.[24] In principle, the exclusion restriction is that $E_{j\tilde{t}}^{-k}$ affects A_{it} only through $A_{j\tilde{t}-1}$. The timings of the two elections across municipalities are highly orthogonal to each other; therefore, the exclusion restriction is likely to be satisfied. The relevance, by design, comes from the strength of the election timing effect of municipality j. Similarly, we instrument the interaction terms ($E_{it}^{-k} \times 1(A_{i\tilde{t}} < A_{j\tilde{t}})$) with the same set of variables interacting with E_{it}^{-k}.

5. Results

5.1. Main results

Table 2 summarizes the main findings of this study. We provide evidence of the election timing effect (our 1st stage) and policy diffusion in a stepwise manner. Finally, we present the results of the interaction between election timing and policy diffusion by fully estimating Equation [1].

Column (1) of Table 2 reports the OLS estimates of the election cycles (α_{-4}, and α_{-1}) *only*—without policy diffusion and its interaction—where the reference year is the middle years. The municipality expands the eligibility age by 0.018 and 0.038 years per month (0.22 and 0.46 years in 12 months) one year *before* the election and four years before the election (or equivalently, one year *after* the previous election), confirming the existence of the election timing effects in the context of child healthcare subsidy in Japan, as graphically seen in Figure 4. By construction, as this effect of election cycles also applies to neighbors j, we next use the election cycle of neighbors as the IV for the neighbors' policy adoption.

[23] Spatially clustered standard errors, as in Conley (1999), do not significantly inflate our standard errors (results available upon request).

[24] We obtain qualitatively similar results without own i's election cycle dummies $E_{i\tilde{t}}^{-k}$ and lagged eligibility ages $A_{i\tilde{t}-1}$ as additional instruments (the results available upon request).

Columns (2) and (3) of Table 2 report the OLS and IV estimates of policy diffusion (β) *only* (without the effect of election timing and their interaction). Here, the weight for each neighbor is based on the level of out-migration, that is, the neighboring municipalities to which their citizens move. The IV estimate in column (3) is smaller than the OLS estimate in column (2), indicating the potential endogeneity of neighbors' policy choices due to positively correlated preferences or environments. Column (3) suggests that the municipality expands the eligibility age by 0.054 years per month (0.65 years in 12 months) when its eligibility age is strictly below that of neighbors, confirming the existence of policy diffusion through geographical channels, as graphically observed in Figure 5. The Kleibergen-Paap-rk Wald-F-statistic is above 80, suggesting that weak identification is unlikely to be a concern in our setting.[25]

Finally, columns (4) and (5) of Table 2 report the OLS and IV estimates of a full Equation [1], which includes the interaction terms of election cycles and policy diffusion. As the neighbors' policy choices seem endogenous, we focus on the IV estimates in column (5). The estimates of the interaction term just before the next election and after the previous election (ρ_{-4} and ρ_{-1}) are positive and highly statistically significant at the 1 percent level, suggesting that municipalities are more likely to adopt a policy to catch up with nearby municipalities around their own election timing. Importantly, the non-interaction terms of policy diffusion (β), which capture policy diffusion in the two middle years, are no longer far from statistically significant, suggesting that policy diffusion occurs only at one's own election timing. While the past studies have examined the determinants of policy diffusion, to the best of our knowledge, we are the first to show that elections play a key role in the processes and mechanisms of policy diffusion: policy diffusion manifests only *through* elections, at least in our setting.

5.2. Heterogeneity

To further understand the underlying mechanism of policy diffusion through elections, we examine how the impact of elections differs according to municipalities' political characteristics. Specifically, we investigate the heterogeneity by electoral competition in the previous election, as well as the heterogeneity by the political experience of mayors. The results are summarized in

[25] To the best of our knowledge, no study has yet developed formal methods for detecting weak identification in the presence of multiple endogenous regressors and heteroskedasticity. As such, we report the Kleibergen-Paap Wald rk F-statistic that is clustered both at own municipality and neighboring municipality level, along with Cragg-Donald F-statistic, which assumes homoskedastic errors.

Table 3.

Columns (1) and (2) of Table 3 show whether policy diffusion through elections is more pronounced after uncontested and contested elections, respectively. Municipalities led by mayors elected through contested previous elections adopt the neighbor's policy around the time of the elections. In particular, the interaction term of four years before the election (i.e., immediately after the previous election) implies that contested elections may force politicians to promise coverage expansion if the municipality's policy is behind that of its neighbors, and they keep the promise after being elected. These results imply that the competitiveness of the previous election may have played a role in policy diffusion.

Columns (3) and (4) of Table 3 examine the heterogeneity in terms of mayors. Column (3) demonstrates that for the 1^{st}-term politicians (that is, elected in the previous election for the first time), the interaction terms are neither statistically significant nor economically large, suggesting that election timing does not influence the timing of the policy adoption by inexperienced politicians. This might reflect the weakness of political foundations of novice politicians to adopt the initiative in decision-making or lack of experience to adopt the policy at the "right" timing. In stark contrast, column (4) demonstrates that the interaction terms for the 2^{nd}+ term politicians are highly statistically significant and larger than the estimates for the full sample. Thus, our main findings on policy diffusion through elections are primarily driven by more experienced politicians who might know the optimal timing of policy adoption, that is, adopting a policy around their election timing to enhance their reelection probability.

5.3. Robustness checks

Which neighbors—. Table 4 reports the estimates from several ways of defining neighborliness: the largest migration outflows (baseline) and inflows in columns (1) and (2), similarity in population size and per capita income in columns (3) and (4). Columns (5) and (6) limit "neighboring" municipalities to the largest and top three to which the citizens move the most. It is reassuring that the estimates, particularly for the interaction terms, are more or less quantitatively similar across different criteria for defining neighborliness.

Time lags—. Thus far, we have arbitrarily chosen three-month lags as the reference period, as discussed in Section 4.1. Table 5 presents the estimates from Equation [1] when the reference period varies from zero to six months, which is equivalent to the intervals between two municipal assemblies. We are reassured that our results are not particularly sensitive to the choice of the reference period.

Robustness—. We subject the main results in column (5) of Table 2 to a series of other robustness checks where "out-migration" is used to define neighborliness and three months as the time lag. Table 6 summarizes the results. For ease of comparison, column (1) replicates our baseline estimates. Column (2) presents the municipalities' linear time trends. However, the estimates are barely affected. Column (3) includes the fixed effects for each of the 12 calendar months for each municipality to account for municipality-specific seasonality. Again, the estimates are similar.

Columns (4)–(6) of Table 6 report the estimates using different sample construction methods. Column (4) excludes the simultaneous election cycles in 2007 and 2011. Column (5) excludes non-scheduled election cycles that were not held four years after previous elections (due to mayors' deaths, for example) during our sample period.[26] Column (6) uses a balanced panel of 221 municipalities (excluding municipal mergers). All the estimates are quantitatively similar to the baseline estimates in column (1). Finally, while our main sample focuses on adjacent municipalities throughout the study, column (7) expands the sample to all municipalities, including non-adjacent municipalities within the same prefecture. While the estimates are slightly smaller than the baseline estimates in column (1), as expected, suggesting that geographical proximity matters for policy diffusion, they are qualitatively similar to the baseline estimates.

Table A3 in the Appendix presents another robustness check. We dropped each prefecture from the sample to determine if the estimates changed drastically. The results showed that none of the prefectures drove the results.

6. Conclusion

This study aimed to understand the determinants of "policy diffusion"—how and why policies spread from jurisdiction to jurisdiction. To explore this question, we collected unique monthly data on subsidies for child healthcare in Japan and exploited the unique institutional setup of exogenously asynchronous election timing in Japan to overcome the identification issue

[26] During our sample period, out of 656 elections, 11.3% (74) had non-scheduled elections due to resignation (36), merger (24), death (7), and others (7).

of estimating policy diffusion.

We document strong evidence that (the timing of) the elections plays a vital role in policy diffusion. Using neighbors' election cycles as an instrument for the neighbors' policy adoption, we find that municipalities are more likely to adopt a policy to catch-up with nearby municipalities around their own election timing. However, we do not find any policy diffusion in the middle two years between elections, suggesting that policy diffusion manifests only through an election. These politicians' behaviors are strategic; not only does such policy diffusion occur only at their own election timings, but municipalities adopt a policy only when their policy falls behind that of a neighbor, and such observations are observed only by experienced politicians.

Notably, we exploit Japan's unique institutional setting in which municipal elections are held at different times to *identify* policy diffusion through elections. Such unsynchronized elections are common. For example, Indian state elections (Khemani 2004; Cole 2009), German local elections across states (Foremny and Riedel 2014; Englmaier and Stowasser 2017), and Italian municipal elections (Repetto 2018) are not synchronized. However, our results may apply to settings such as the US states, where elections are held simultaneously. In fact, a state may similarly adopt a policy at election timing by comparing its current policy with that of neighboring states, even though simultaneous elections do not provide researchers with variations in election timing to identify policy diffusion through elections.

References

Akhmedov, A., Zhuravskaya, E. (2004) "Opportunistic political cycles: test in a young democracy setting." *Quarterly Journal of Economics*, 119(4): 1301–1338.

Arellano, M., Bond, S. (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58(2): 277–297.

Baicker, K. (2005) "The spillover effects of state spending." *Journal of Public Economics*, 89(2‐3): 529–544.

Banerjee, A.V. (1992) "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107(3): 797–817.

Besley, T, Case, A. (1995) "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition." *American Economic Review*, 85(1): 25–45.

Brender, A., Drazen, A. (2005) "PBC in new versus established democracies." *Journal of Monetary Economics*, 52(7): 1271–1295.

Case, A.C., Hines Jr., J.R., Rosen, H.S. (1993) "Budget spillovers and fiscal policy interdependence: evidence from the States." *Journal of Public Economics*, 52: 285–307.

Cole, S. (2009) "Fixing Market Failures or Fixing Elections? Agricultural Credit in India." *American Economic Journal: Applied Economics*, 1(1): 219–250.

Conley, T.G. (1999) "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

de Haan, J., Klomp, J. (2013) "Conditional PBC: a review of recent evidence." *Public Choice*, 157: 387–410.

de Paula, A., Rasul, I., Souza, P. CL. (2020) "Identifying Network Ties from Panel Data: Theory and an Application to Tax Competition." *Working paper*.

DellaVigna, S., Kim, W. (2022) "Policy Diffusion and Polarization across U.S. States." *NBER Working Paper No. 30142.*

Drazen, A. (2001) "The political business cycle after 25 years." In: Bernanke, B.S., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2000, 75−117. MIT Press, Cambridge, MA.

Englmaier, F., Stowasser, T. (2017) "Electoral Cycles in Savings Bank Lending." *Journal of the European Economic Association*, 15(2): 296–354.

Foremny, D., Riedel, N. (2014) "Business taxes and the electoral cycle." *Journal of Public Economics*, 115: 48–61.

Fukumoto, K., Ueki, F. (2015) "Shichoson senkyo ga toitsu chiho senkyo kara itsudatsu shita jiki to riyu." (In Japanese) [When and why did municipalities drop from the simultaneous local elections?]. *Gekkan Senkyo*, 9–11: 8–15, 9–14, 17–24.

Fukumoto, K., Horiuchi, Y. (2011) "Making outsiders" votes count: Detecting electoral fraud through a natural experiment." *American Political Science Review*, 105(3): 586–603.

Fukumoto, K., Horiuchi, Y. (2016) "Identifying the effect of mobilization on voter turnout through a natural experiment." *Electoral Studies*, 44: 192–202.

Gibbons, S., Overman, H. G. (2012) "Mostly Pointless Spatial Econometrics?" *Journal of Regional Science*, 52(2): 172–191.

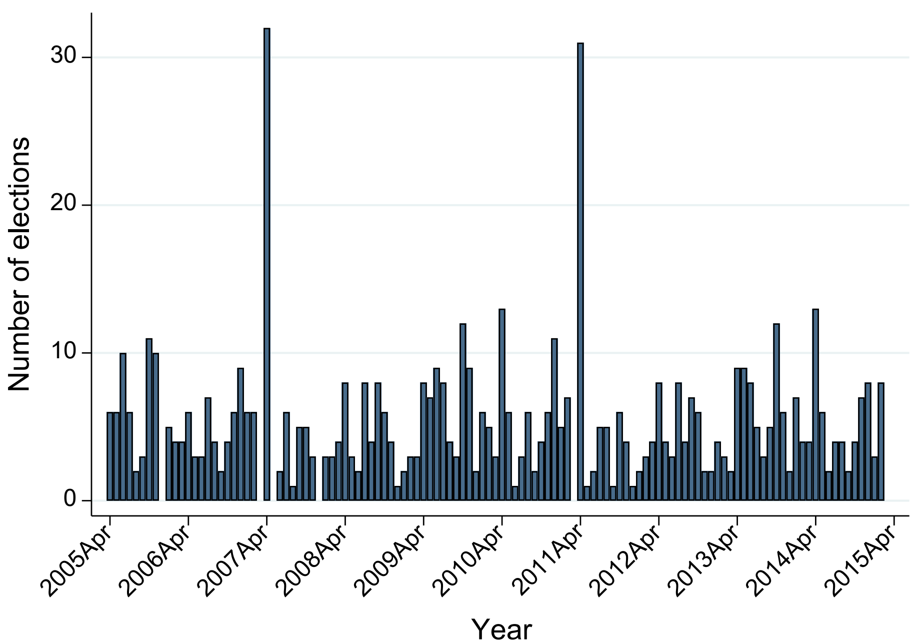
Iizuka, T., Shigeoka, H. (2018) "Free for Children? Patient Cost-sharing and Health Care Utilization." *NBER Working Paper No. 25306.*

Iizuka, T., Shigeoka, H. (2022a) "Is Zero a Special Price? Evidence from Child Health Care." *American Economic Journal: Applied Economics*, 14(4): 381–410.

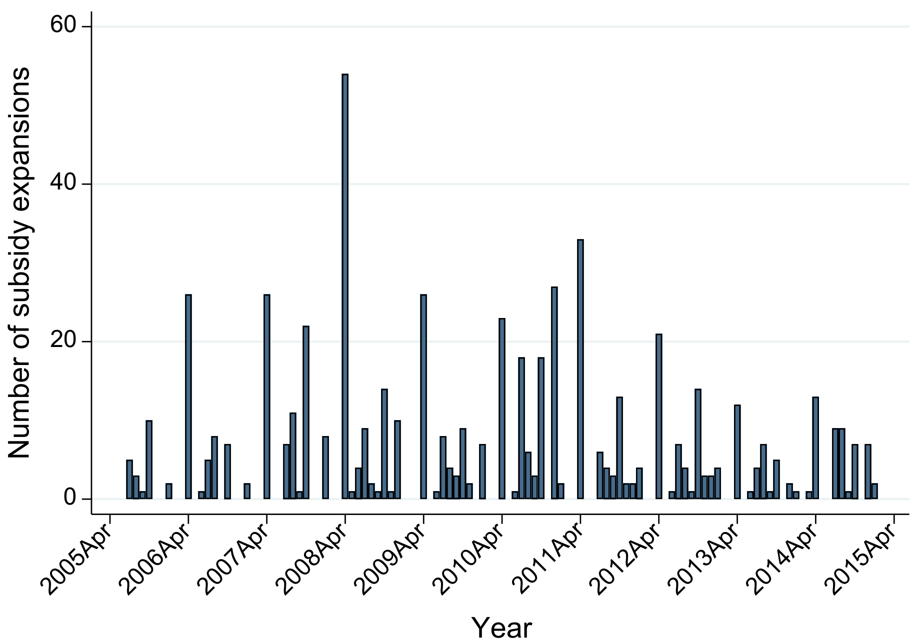
Iizuka, T., Shigeoka, H. (2022b) "Asymmetric Demand Response when Prices Increase and Decrease: The Case of Child Healthcare." *The Review of Economics and Statistics*, forthcoming.
Ikegami, N., Campbell, J. C. (1995) "Medical care in Japan." *New England Journal of Medicine*, 333: 1295–1299.
Khemani, S. (2004) "Political cycles in a developing economy: effect of elections in the Indian states." *Journal of Development Economics*, 73(1): 125–154.
Kondo, A., Shigeoka, H. (2013) "Effects of universal health insurance on health care utilization and supply-side responses: Evidence from Japan." *Journal of Public Economics*, 99: 1–23.
Manski, C. (1993) "Identification of endogenous social effects: The reflection problem." *The Review of Economic Studies*, 60(3): 531–542.
Mallinson, D. J. (2021) "Who are your neighbors? The role of ideology and decline of geographic proximity in the diffusion of policy innovations." *Policy Studies Journal*, 49(1): 67–88.
Nordhaus, W. (1975) "The political business cycle." *Review of Economic Studies*, 42: 169–190.
Repetto, L. (2018) "Political Budget Cycles with Informed Voters: Evidence from Italy." *Economic Journal,* 128: 3320–3353.
Rogoff, K. (1990) "Equilibrium PBC." *American Economic Review*, 80: 21–36.
Rogoff, K., Sibert, A. (1988) "Elections and macroeconomic policy cycles." *Review of Economic Studies*, 55: 1–16.
Shigeoka, H. (2015) "School entry cutoff date and the timing of births." *NBER Working Paper No. 21402*.
Volden, C., Ting, M., Carpenter, D. (2008) "A formal model of learning and policy diffusion." *American Political Science Review*, 102(3): 319–332.
Walker, J. L. (1969) "The Diffusion of Innovations among the American States." *American Political Science Review*, 63(3): 880–899.

Figure 1: Exact dates of elections and subsidy expansions

Panel A. Municipalities holding elections (monthly)

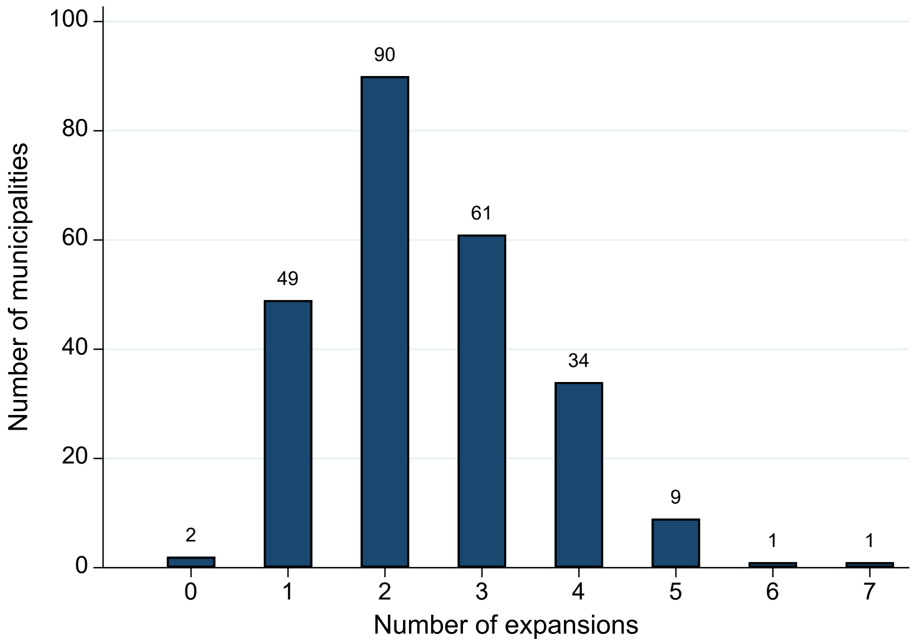


Panel B. Municipalities experiencing subsidy expansions (monthly)



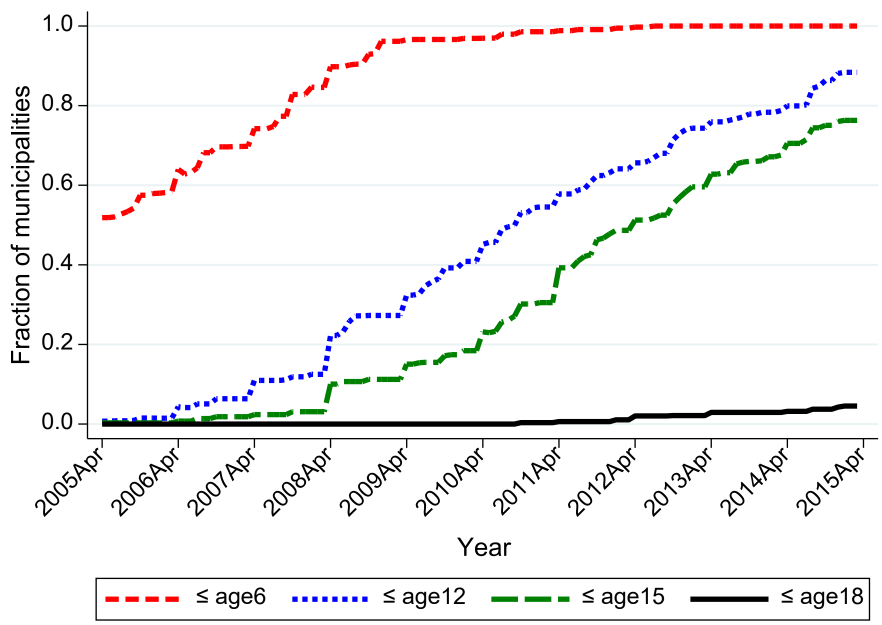
Notes: Panel A plots the number of municipalities that held elections each month between April 2005 and March 2015. There were 632 elections in total. Panel B shows the number of municipalities that experienced subsidy expansion each month during the same period. A total of 606 subsidy expansions occurred. The total number of municipalities was 247.

Figure 2: Number of subsidy expansions



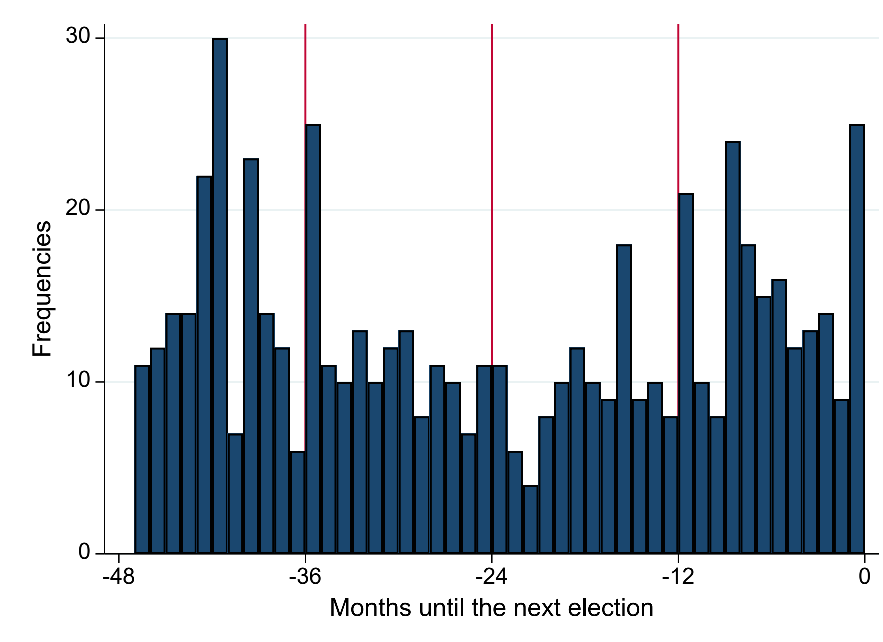
Notes: The figure plots the number of municipalities that experienced a particular number of subsidy expansions from April 2005 to March 2015 (see Panel B of Figure 1 for the precise timing of all policy changes). Only two out of 247 municipalities did not experience any subsidy expansion. A total of 606 subsidy expansions occurred. The average number of expansions per municipality was 2.45, as listed in Table 1.

Figure 3: Time series of maximum age covered by healthcare subsidy



Notes: The figure plots the share of municipalities in our insurance claims data by the maximum age for subsidy eligibility for outpatient care at the monthly level from April 2005 to March 2015 (see Figure 1-B for the precise timing of all policy changes). There were a total of 247 municipalities. The ages of 6, 12, 15, and 18 correspond to entering elementary school, graduation from elementary school, graduation from junior high school, and graduation from high school, respectively.

Figure 4: Timing of the subsidy expansions vis-à-vis election



Notes: The figure plots the number of subsidy expansions in the months until the next election. A total of 606 subsidy expansions occurred. There were a total of 247 municipalities.

Figure 5: The official gazette for elections

平成30年4月15日執行
津島市長選挙
選挙公報
津島市選挙管理委員会

中3まで医療費無料を！

医療費完全無料化をめざし、その実現に向けて市長給与70%削減と退職金1,812万円の全額返還、これを財源の一部に充てます。

無所属
杉山 良介
すぎやま りょうすけ
78歳

☆子どもは、津島の宝です。

その為に市長の給料は、
70%カット
年間約1,514万円の70%
＝1,060万円カット
4年間で1,060万円×4
＝4,240万円カット
退職金1,812万円を全額返還　合計で、4,240万円＋1,812万円＝6,052万円
これを医療費無料化の財源の一部にあてます。

市議36年の経験生かし、
福祉・くらしを守る市政を

いますぐ実現を！

☆中3まで医療費無料
☆国保、介護の負担軽減を
☆地元業者応援で地域を元気に

つしま成長戦略 第2弾
3つのKEY WORD 透明 共に 地域
子どもたちが津島で夢がかなえられる、まちづくり

❶ 子ども・子育て応援都市、つしま
子どもが生まれる前から、産み育てるまで、丸ごと応援します
● 中学卒業までの子ども医療費を無料。 ● 子どもが学び育つ教育環境の整備。 ● 丸ごと応援する子育て世代包括支援センターの活動を充実。

❷ 防災・減災のモデル都市、つしま
海抜ゼロメートル地帯のまちづくりモデルを、津島でつくります
● 市民全員を対象に防災教育、防災学習の機会を提供。 ● 災害時に地域で助け合うための仕組みづくり。 ● 災害時に備えて、県境を越えた自治体との関係づくり。

❸ 地域の特性を活かした交流都市、つしま
津島神社を核として、年間を通じて歴史や文化が体感できるゾーンづくりを進めます
● ユネスコ無形文化遺産登録を機に「津島―Tsushima」を海外に発信。 ● 祭りを核とした広域観光・交流。 ● 歴史的風致維持向上計画を策定し、歴史・文化のまちづくりを推進。

❹ 地域経済が活性化する発展都市、つしま
元気な津島へ。市内で「働く場」の確保に取り組みます
● 企業誘致や、起業環境を整備。 ● 津島駅西エリアの新生や駅を活かしたまちづくり。
● 市民参加で市内全域の都市計画の見直しを推進。

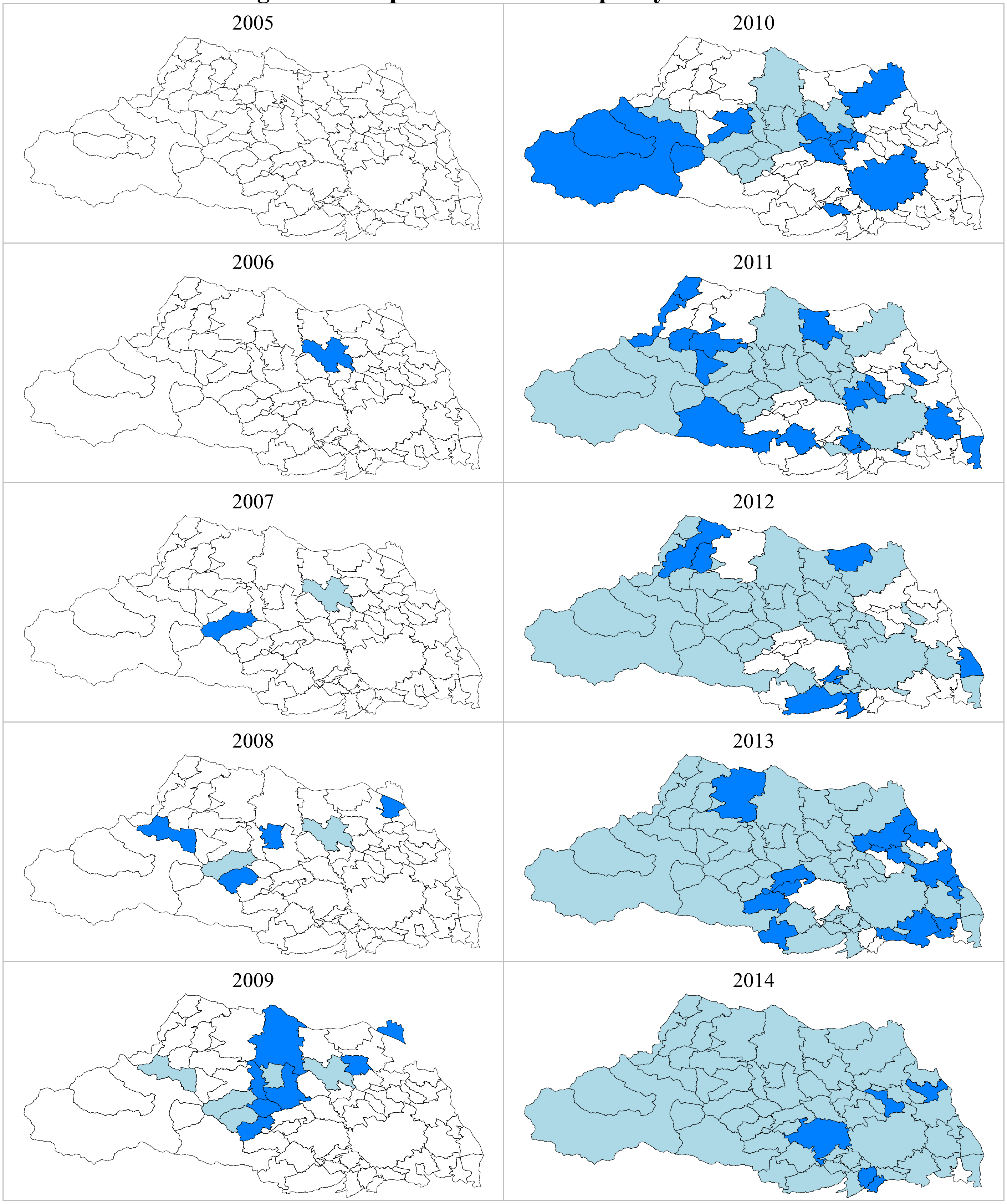
❺ いつまでも健康で暮らす都市、つしま
「健康のまち津島」を継承し、すべての人にやさしいまちづくりに取り組みます
● ニーズに応える医療提供と病院経営を推進。 ● すべての人を対象とする地域包括ケアシステムを構築。 ● 地域のコミュニティなどと一緒に、健康教育・健康教室を開催。

市政持続のために ひと・もの・ことの改革
市役所職員の能力を高めるとともに、定員管理と組織改革をすすめます。現在、市の財政は厳しい状況にあります。人口も減っています。過去に建設された公共施設が更新時期を迎えています。学校を核とした施設の再編、統廃合をすすめます。地域の資源と力を活かして、地域創造を計画的にすすめていきます。ひと・もの・ことの改革をすすめるにあたり、外部委員会を設けて有識者からの提言・助言をいただくとともに、タウンミーティングなどにより広く市民の皆様からのご意見をいただきすすめていきます。

ひび 一昭
かずあき
無所属
65歳

Notes: The figure shows the official gazette for the municipal election in Tsushima City in Aichi Prefecture, held on April 15, 2018. The sentences in the red box indicate subsidy expansion for child healthcare. The candidate on the right is the incumbent (ひび 一昭), who promises to raise the eligibility age for free healthcare till the end of junior high school (中学卒業 on the right or 中 3 on the left in the gazette), which is equivalent to age 15. The candidate on the left is the opponent (杉山 良介), who also promises the same subsidy expansion. The incumbent won this election and implemented the policy one year later on April 1, 2019.

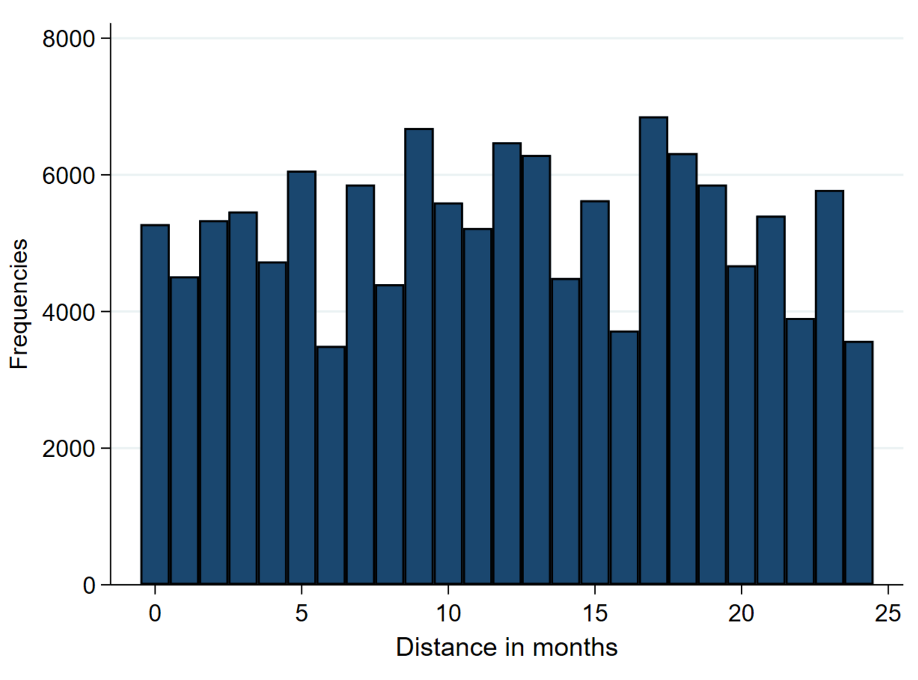
Figure 6: Graphical evidence of policy diffusion



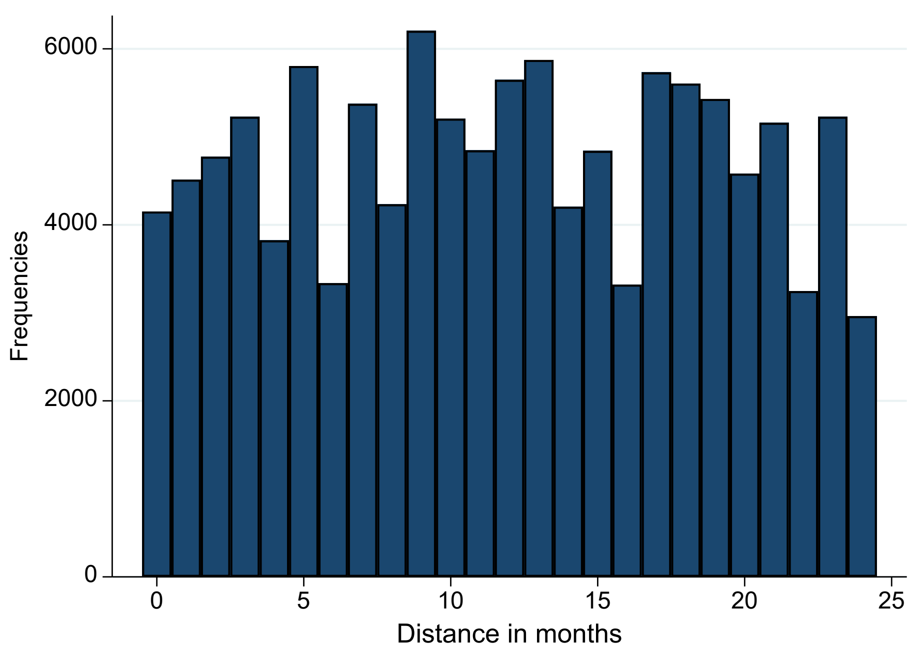
Notes: Each graph describes the subsidy level every April from 2005 to 2014 in Saitama Prefecture. Darker colors indicate that the municipalities have expanded the subsidy to age 15 (the end of junior high school) in the year. The lighter color indicates that municipalities have expanded their subsidies to the age of 15 in the past.

Figure 7: Orthogonality of elections among neighbors

Panel A. Distance between the elections of adjacent municipalities



Panel B. Distance between the elections of adjacent municipalities (excluding simultaneous elections)



Notes: Panel A plots the distance (in months) between the election timings of adjacent municipalities from April 2005 to March 2015 and Panel B plots the same figure after excluding simultaneous elections. Given that one election cycle lasts 48 months, the distance between two election cycles cannot exceed 24 months.

Table 1: Summary statistics

Variable	N	Mean	SD	Min	Max
A. Subsidy characteristics					
Expansion dummy	127,288	0.02	0.14	0	1
Eligibility age (A_i)	127,288	9.48	4.33	2.5	18
No more than 6 ($A_i \leq 6$)	127,288	0.89	0.31	0	1
No more than 9 ($A_i \leq 9$)	127,288	0.51	0.50	0	1
No more than 12 ($A_i \leq 12$)	127,288	0.42	0.49	0	1
No more than 15 ($A_i \leq 15$)	127,288	0.29	0.46	0	1
No more than 18 ($A_i \leq 18$)	127,288	0.01	0.11	0	1
1 ($A_i < A_j$)	127,288	0.23	0.42	0	1
B. Election characteristics					
Female	2,684	0.02	0.13	0	1
Terms	2,684	2.12	1.26	1	10
1^{st} term	2,684	0.39	0.49	0	1
2^{nd}+ term	2,684	0.61	0.49	0	1
Uncontested election	2,684	0.19	0.39	0	1
Scheduled election	2,684	0.88	0.32	0	1
Simultaneous election	2,684	0.06	0.23	0	1
C. Municipality characteristics					
Population between 0-14	127,288	0.14	0.02	0.08	0.19
Population between 15-64	127,288	0.65	0.04	0.44	0.75
Population between 65+	127,288	0.21	0.05	0.09	0.48
Population density	127,288	2,705	2,691	9	14,020
Per capita income	127,288	3.25	0.39	2.41	4.94

Notes: Subscripts i and j indicate own and neighboring municipalities, respectively. Note that we allowed multiple observations of neighboring municipalities (j) per municipality (i). Panels A and B were manually collected by the authors. For Panel C, all variables were obtained from "Sichoson no Sugata," published by the Statistics Bureau (https://www.e-stat.go.jp/regional-statistics/ssdsview, last accessed on August 1, 2019).

Table 2: Main results

	Election cycle	Policy diffusion		With interactions	
	OLS	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)
1 year before election	0.018**			0.007	-0.005
	(0.008)			(0.007)	(0.008)
4 years before election	0.038***			0.029***	0.014*
	(0.008)			(0.007)	(0.007)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$		0.077***	0.054***	0.055***	0.003
		(0.009)	(0.011)	(0.011)	(0.015)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$				0.046**	0.097***
				(0.021)	(0.036)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$				0.045**	0.110***
				(0.020)	(0.034)
R-squared	0.98	0.98	0.89	0.98	0.89
N	127,288	127,288	127,288	127,288	127,288
N of municipalities	247	247	247	247	247
N of neighing municipalities (average)	-	4.47	4.47	4.47	4.47
Cragg-Donald-Wald F-statistic			8,120.0		2,401.9
Kleibergen-Paap-rk Wald-F-statistic			81.9		29.0

Notes: The outcome is the eligibility age for subsidies. Column (1) reports the OLS estimates of the election cycles (α_{-k} $(k = 1, 4)$) only from Equation [1], and columns (2) and (3) report the OLS and IV estimates of policy diffusion (β) only from Equation [1]. Here, the neighbor is chosen from among the "neighboring" municipalities to which their citizens move the most. Columns (4) and (5) report the estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1], with standard errors clustered at both municipality i and neighbor municipality j levels reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 3: Heterogeneity

	Type of previous election		Terms	
	Uncontested	Contested	1st term	2nd+ term
	(1)	(2)	(3)	(4)
1 year before election	0.023	-0.013	-0.016	-0.011
	(0.026)	(0.009)	(0.016)	(0.012)
4 years before election	0.042*	0.018**	0.045***	0.023**
	(0.023)	(0.009)	(0.017)	(0.011)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.069*	-0.019	0.039	0.007
	(0.036)	(0.016)	(0.026)	(0.021)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.142	0.109***	0.070	0.152***
	(0.102)	(0.038)	(0.056)	(0.054)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	-0.088	0.128***	-0.036	0.112**
	(0.070)	(0.035)	(0.052)	(0.049)
R-squared	0.84	0.89	0.85	0.88
N	22,063	105,224	49,836	77,451
Cragg-Donald-Wald F-statistic	456.7	2,018.3	905.4	1,592.3
Kleibergen-Paap-rk Wald-F-statistic	11.1	27.6	16.7	26.1

Notes: The outcome is an eligibility age for the subsidy. The estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1] are reported with standard errors clustered at both municipality i and neighbor municipality j levels reported in parentheses. “Out-migration,” which determines the degree of neighborliness by the fraction of those that move into each neighboring municipality, is used to construct the neighborliness. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 4: Choice of neighbors

Choice of neighbor	*Out-migration (baseline)*	*In-migration*	*Size of population*	*Per capita income*	*Out-migration (top1)*	*Out-migration (top3)*
	(1)	(2)	(3)	(4)	(5)	(6)
1 year before election	-0.005	-0.006	-0.002	-0.007	-0.007	-0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
4 years before election	0.014*	0.012	0.012	0.010	0.016**	0.015**
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.001	-0.001	-0.010	-0.010	0.048**	0.005
	(0.016)	(0.015)	(0.017)	(0.016)	(0.023)	(0.016)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.096***	0.101***	0.082**	0.099***	0.106**	0.098***
	(0.035)	(0.035)	(0.032)	(0.035)	(0.041)	(0.037)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.111***	0.119***	0.115***	0.123***	0.108***	0.108***
	(0.034)	(0.034)	(0.031)	(0.033)	(0.038)	(0.035)
R-squared	0.89	0.89	0.90	0.90	0.90	0.90
N	127,288	127,288	127,288	127,288	28,454	80,322
Cragg-Donald-Wald F-statistic	2,405.2	2,403.5	2,563.0	2,515.0	390.1	1,431.7
Kleibergen-Paap-rk Wald-F-statistic	29.2	28.4	28.7	30.5	18.3	26.7

Notes: The outcome is the eligibility age for subsidies. The estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1] are reported, with standard errors clustered at both municipality i and neighboring municipality j levels reported in parentheses. Column (1) replicates column (5) of Table 2 for ease of comparison, which determines the degree of neighborliness by the fraction of those that move into each neighboring municipality (out-migration). "In-migration" in column (2) determines the degree of neighborliness by the fraction of those that move from (in-migration) each neighboring municipality. "Size of the population" and "per capita income" in columns (3) and (4) compute weight on the difference in population and per capita income, reflecting that neighbors with similar size of population or per capita income receive more weight. Finally, columns (5) and (6) limit the neighboring municipalities to the largest and top three to which citizens move the most. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 5: Length of lag

t is lagged by x months	*0*	*1*	*2*	***3***	*4*	*5*	*6*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 year before election	-0.003	-0.004	-0.005	**-0.005**	-0.002	-0.004	-0.001
	(0.008)	(0.008)	(0.008)	**(0.008)**	(0.008)	(0.008)	(0.008)
4 years before election	0.018**	0.017**	0.015**	**0.014***	0.015**	0.013*	0.011
	(0.007)	(0.007)	(0.007)	**(0.007)**	(0.007)	(0.008)	(0.008)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.001	0.002	0.008	**0.002**	0.000	0.008	-0.004
	(0.015)	(0.016)	(0.015)	**(0.015)**	(0.015)	(0.015)	(0.016)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.093**	0.095**	0.099***	**0.097*****	0.087**	0.088**	0.079**
	(0.038)	(0.037)	(0.037)	**(0.035)**	(0.035)	(0.035)	(0.033)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.096***	0.098***	0.106***	**0.112*****	0.103***	0.109***	0.106***
	(0.034)	(0.034)	(0.034)	**(0.034)**	(0.033)	(0.033)	(0.033)
R-squared	0.90	0.90	0.90	**0.90**	0.89	0.90	0.90
N	130,657	129,534	128,410	**127,288**	126164	125,042	123,918
Cragg-Donald-Wald F-statistic	3,008.0	2,264.0	2,305.5	**2,401.3**	2,505.4	2,575.0	2,629.6
Kleibergen-Paap-rk Wald-F-statistic	31.3	26.5	28.3	**29.3**	29.5	29.5	29.1

Notes: The outcome is the eligibility age for subsidies. The estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1] are reported, with standard errors clustered at both municipality i and neighboring municipality j levels reported in parentheses. "Out-migration," which determines the degree of neighborliness by the fraction of those that move into each neighboring municipality, is used to construct the neighborliness. Column (4), with a three-month lag, is the baseline, which is identical to column (5) of Table 2. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 6: Robustness checks

	Baseline	Mun trend	Each calendar month FE. by mun	Drop simultaneous elections	Drop non-scheduled elections	Balanced panel	All municipalities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 year before election	-0.005	-0.006	-0.002	-0.006	-0.001	-0.003	-0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
4 years before election	0.014*	0.015*	0.016**	0.014*	0.012	0.013	0.015**
	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.002	0.020	0.010	-0.003	-0.002	0.004	-0.012
	(0.015)	(0.016)	(0.015)	(0.017)	(0.017)	(0.016)	(0.012)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.096***	0.088**	0.090***	0.120***	0.099**	0.099***	0.079**
	(0.035)	(0.037)	(0.034)	(0.043)	(0.039)	(0.036)	(0.031)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.111***	0.106***	0.104***	0.119***	0.119***	0.101***	0.098***
	(0.034)	(0.034)	(0.032)	(0.038)	(0.039)	(0.033)	(0.030)
R-squared	0.90	0.85	0.90	0.89	0.89	0.89	0.89
N	127,288	127,288	127288	110,628	111,590	120,072	1,235,586
Cragg-Donald-Wald F-statistic	2,402.3	2,532.7	2,413.1	2,050.2	2,258.2	2,304.1	24,823.1
Kleibergen-Paap-rk Wald-F-statistic	29.2	30.4	29.3	25.0	26.0	27.7	28.9
Mun FE, Prefecture-year-month FE	X	X	X	X	X	X	X
Other covariates	X	X	X	X	X	X	X
Mun trend		X					
Calendar month by mun FE			X				

Notes: The outcome is the eligibility age for subsidies. The estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1] are reported, with standard errors clustered at both municipality i and neighboring municipality j levels reported in parentheses. "Out-migration," which determines the degree of neighborliness by the fraction of those that move into each neighboring municipality, is used to construct the neighborliness. Column (1) replicates the baseline estimates from column (5) of Table 2. Column (2) shows the municipality-specific linear trends. Column (3) includes fixed effects (FE) for each of the 12 calendar months in each municipality to control for municipality-specific seasonality. Column (4) excludes the simultaneous election cycles. Column (5) excludes non-scheduled election cycles. Column (6) uses a balanced panel of 221 municipalities. Column (7) uses all the municipalities, including non-adjacent ones. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

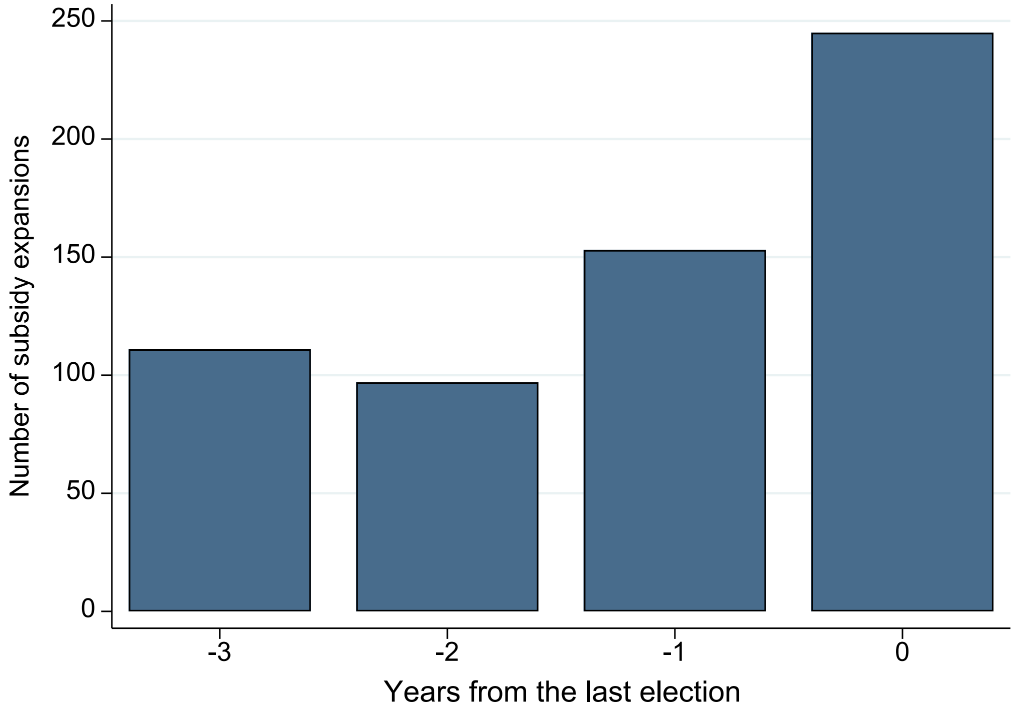
Online Appendix

(Not for Publication)

Policy Diffusion Through Election

by Hitoshi Shigeoka and Yasutora Watanabe

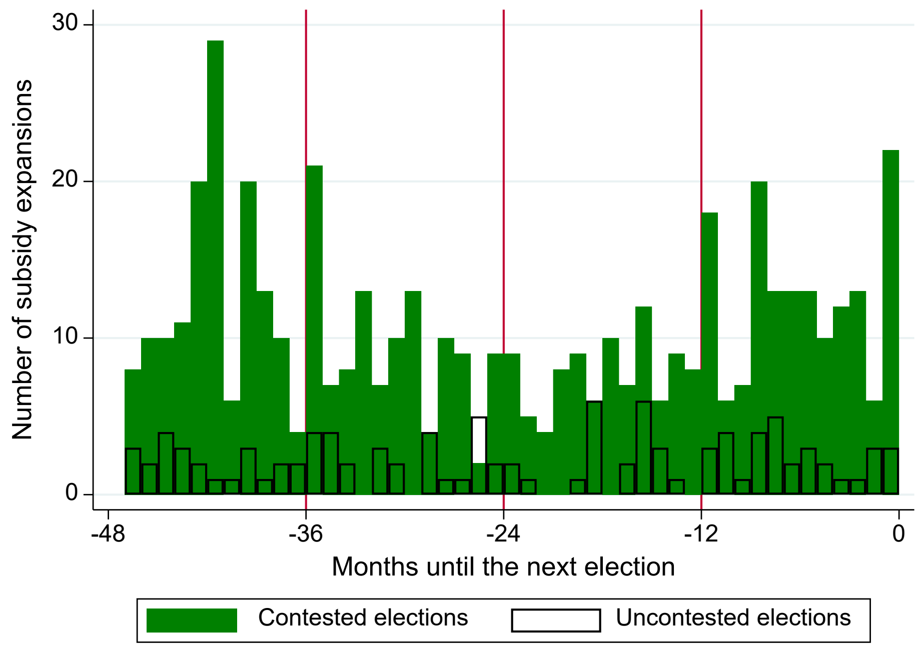
Figure A1: Year-level aggregations



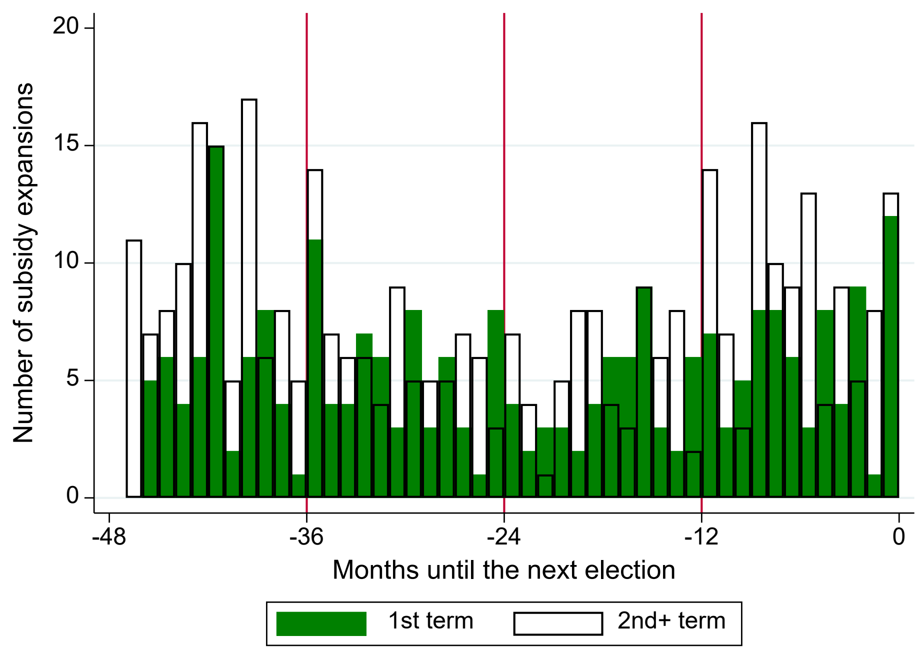
Notes: This figure illustrates the number of subsidy expansions by year until the next election, assuming that we have only yearly information on when subsidy expansion was implemented.

Figure A2: Timing of the subsidy expansions (heterogeneity)

Panel A. Contested vs. uncontested elections



Panel B. 1st term vs. 2nd+ term



Notes: Panel A plots the number of subsidy expansions by the month until the next election for two types of elections: contested and uncontested in the *previous* elections. Of the 606 subsidy expansions, 500 (82.5%) were contested and 106 (17.5%) were uncontested. Panel B plots the same for the 1st term and 2nd+ term during the current election cycle (i.e., -48 months to 0). 245 (40.4%) were implemented during the first term, and 361 (59.6%) were implemented during the 2nd+ term.

Table A1: Balanced checks

	Simultaneous elections	Not in simultaneous elections	*Dif* =(1)-(2)
	(1)	(2)	(3)
A. 2007 elections			
Population between 0–14	0.14	0.14	0.00
	[0.01]	[0.02]	(0.00)
Population between 15–64	0.67	0.67	0.00
	[0.05]	[0.04]	(0.01)
Population at 65 or above	0.19	0.19	0.00
	[0.06]	[0.05]	(0.01)
Population density	3648.60	2535.45	762.22
	[2851.94]	[2614.92]	(483.51)
Per capita income	3.47	3.40	0.01
	[0.34]	[0.41]	(0.07)
Number of municipalities	32	214	
B. 2011 elections			
Population between 0–14	0.13	0.13	0.00
	[0.02]	[0.02]	(0.00)
Population between 15–64	0.63	0.64	0.00
	[0.04]	[0.04]	(0.01)
Population at 65 or above	0.23	0.23	0.01
	[0.06]	[0.05]	(0.01)
Population density	3629.95	2589.73	653.15
	[2870.76]	[2681.58]	(506.50)
Per capita income	3.21	3.15	0.01
	[0.33]	[0.36]	(0.07)
Number of municipalities	31	216	

Notes: The table compares the municipal characteristics across the two groups (with and without simultaneous elections) in the 2007 (Panel A) and 2011 (Panel B) elections. We include prefecture FE when comparing columns (1) and (2); therefore, column (3) does not simply match the difference between columns (1) and (2).

Table A2: Distribution of eligibility age (A_i)

A_i	N	%
2.5	1,283	4.36
3.5	1,308	4.44
4.5	709	2.41
5	12	0.04
5.5	360	1.22
6	10,301	35.00
6.5	291	0.99
7	353	1.20
7.5	24	0.08
8	105	0.36
9	2,527	8.59
9.5	24	0.08
10	180	0.61
11	36	0.12
12	3,548	12.06
15	7,957	27.04
16	32	0.11
17	24	0.08
18	354	1.20
Total	29,428	100

Notes: The unit of observation is municipality-year-time. The ages of 6, 12, 15, and 18 correspond to entering elementary school, graduation from elementary school, graduation from junior high school, and graduation from high school, respectively. The age of 9 years corresponds to the 3rd grade of elementary school. Ages 6, 9, 12, and 15 years accounted for 82.7% of all age distributions. Only 1.39% were above the age of 15 years, indicating ceiling effects.

Table A3: Drop one prefecture at a time

Exclude	*Saitama*	*Chiba*	*Kanagawa*	*Aichi*	*Osaka*
	(1)	(2)	(3)	(4)	(5)
1 year before election	0.002	-0.007	-0.005	-0.008	-0.005
	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)
4 years before election	0.009	0.023***	0.013	0.012	0.015*
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
$1(A_{i\tilde{t}} < A_{j\tilde{t}})$	-0.006	-0.005	0.005	0.007	0.013
	(0.016)	(0.018)	(0.017)	(0.016)	(0.017)
1 year before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.081**	0.099**	0.107**	0.087**	0.109***
	(0.034)	(0.043)	(0.041)	(0.040)	(0.040)
4 years before election × $1(A_{i\tilde{t}} < A_{j\tilde{t}})$	0.112***	0.093**	0.130***	0.083**	0.135***
	(0.035)	(0.039)	(0.038)	(0.036)	(0.038)
R-squared	0.89	0.90	0.89	0.90	0.89
N	91,452	100,886	110,776	98,702	107,336
Cragg-Donald-Wald F-statistic	1,751.1	1,834.8	2,099.7	1,970.0	1959.7
Kleibergen-Paap-rk Wald-F-statistic	28.8	23.7	24.9	24.9	26.2

Notes: The outcome is the eligibility age for subsidies. The estimates α_{-k} $(k = 1, 4)$, β, and their interactions ρ_{-k} $(k = 1, 4)$ from Equation [1] are reported, with standard errors clustered at both municipality i and neighboring municipality j levels reported in parentheses. Each column reports the estimates obtained by dropping one prefecture at a time. Significance levels: *** p<0.01, ** p<0.05, * p<0.10