RIETI Discussion Paper Series 20-E-036

Heterogeneous Treatment Effects of Place-based Policies: Which Cities Should be Targeted?

FUJISHIMA, Shota
Hitotsubashi University

HOSHINO, Tadao
Waseda University

SUGAWARA, Shinya
Tokyo University of Science

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 20-E-036

April 2020

Heterogeneous treatment effects of place-based policies: Which cities should be targeted?*

Shota Fujishima¶

Tadao Hoshino§

Shinya Sugawara#

Abstract

In this paper, we empirically assess the treatment effects of a Japanese place-based policy known as the urban revitalization zone (URZ) program. on regional economies. We propose socially desirable assignment rules for the URZ program. The mixed results of previous empirical studies on place-based policies suggest that the treatment effects are regionally heterogeneous. In order to account for such heterogeneity across regions, this study estimates the conditional average treatment effect (CATE) for each region using a marginal treatment effects framework. We then use the estimated CATE parameters to construct empirically welfare-maximizing treatment rules based on the following three types of regional characteristic variables: demographic variables, suburbanization variables, and local production-network variables. Our results indicate that the treatment choice rule based on the suburbanization variables is the most successful. In particular, we find that cities with high numbers of cars per household and low percentages of large-scale stores should be targeted for URZ.

Keywords: conditional average treatment effect, place-based policy, treatment choice.
JEL classification: C15, C31, C33, C36, R58

* This study is conducted as a part of the Project "An Empirical Study on Compact Cities: Evaluating place-based policies in Japan" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). We are grateful to Adrian Lerche and Ryo Nakajima for their helpful comments and discussions. We also thank the seminar audience at Hitotsubashi University, Kyoto University, The Research Institute of Economy, Trade and Industry (RIETI), and the University of Tokyo, and participants at the 14th Meeting of the Urban Economics Association and the 30th Annual Meeting of the Applied Regional Science Conference. Any remaining errors are our own.

¶ Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. Phone/Fax: +81-42-580-8854, Email: s.fujishima@r.hit-u.ac.jp

§ School of Political Science and Economics, Waseda University, 1-6-1 Nishi-waseda, Shinjuku, Tokyo 169-8050, Japan. Email: thoshino@waseda.jp

School of Management, Tokyo University of Science, 1-11-2, Fujimi, Chiyoda, Tokyo 102-0071, Japan. Email: shinya_sugawara@rs.tus.ac.jp

1 Introduction

In this paper, we estimate the causal impacts of a place-based policy in Japan, the urban revitalization zone (URZ) program, on regional economic outcomes. The program consists of the incentives provided by the Act on Vitalization in City Center. Place-based policies such as the URZ program in Japan are introduced to enhance economic performance of particular underperforming cities or regions. Although the Japanese URZ program provides several types of incentives, the primary one is the scheme that encourages the development of commercial facilities in URZ, where, in addition to the subsidies to local agents such as urban development companies that develop commercial facilities, firms that do their business in those facilities can receive financial assistance through preferential interest rates.[1] Neumark and Simpson (2015) summarize in their Table 18.1 various types of place-based policies that have been implemented in various countries.

When the government implements a place-based policy, it needs to decide which places should be targeted with higher priority based on some criterion. However, because the government generally faces a severe budget constraint, the number of places that can enjoy the policy is limited. Hence, in which city the policy maker should implement the policy is an extremely important policy question. To address this type of policy question, the literature on optimal (in some sense) *treatment choice* is rapidly growing (see, notably, Manski, 2004; Bhattacharya and Dupas, 2012; Kitagawa and Tetenov, 2018). Thus, following these studies, if we can propose an optimal treatment choice rule for place-based policy, it will be a useful guideline for the policy makers.

Clearly, if the effects of place-based policy were homogeneous across cities, the optimal treatment assignment rule would be trivial; any assignment rules do not make any significant differences as long as the numbers of the cities targted by the policy are the same. However, this assumption may not be realistic. Indeed, there is a large empirical literature on place-based policies, and they have shown mixed results regarding the effectiveness of such policies (see, e.g., Neumark and Simpson (2015) for an extensive survey).[2] This would imply that the program effects may vary with regional-specific social and economic conditions, as pointed out by Kline and Moretti (2014).

The heterogeneity of treatment effects of place-based policies in terms of regional characteristics has not been well explored, with a few exceptions. Kolko and Neumark (2010) study how the effects of California's enterprise zone (EZ) program vary with zone characteristics such as the share of manufacturing employment and activities that local administrators engage in to achieve the program's goal. The authors collect the latter information by conducting interviews with local administrators. They found that the program is likely to have positive effects on job growth in zones where the share of manufacturing employment is low and local administrators direct their efforts toward marketing and outreach activities. Briant et al. (2015) estimate treatment effects of a place-based policy in France called the Zones Franches Urbanies. They evaluate the policy for each of spatially-integrated neighborhoods and spatially-isolated neighborhoods in terms of attributes such as road access and found that, whereas the policy has no significant effect on wage in spatially-integrated neighborhoods, it has

[1] We introduce the program in more detail in Section 2.

[2] More recent contributions include Chaurey (2016), Ehrlich and Seidel (2018), Givord et al. (2018), Lu et al. (2019), and Neumark and Young (2019).

a positive effect on wage in spatially-isolated neighborhoods.[3]

The two aforementioned papers have employed regression-based difference-in-differences (DID) methods as their empirical strategy.[4] As is well known, under the standard DID assumptions, the causal effect that can be identified by DID is the average treatment effect on the treated (ATET). However, in order to propose an optimal treatment assignment rule, the parameter we need to know is the average treatment effect conditional on the regional characteristics (CATE: conditional average treatment effect) for all data observations rather than the one restricted only to the treated observations. Here, note that standard techniques used to recover the CATE parameter, such as matching and propensity score methods, are based on some conditional independence assumptions on the treatment variable. However, the treatment variable in our case (i.e., the designation of the URZ program) should be naturally treated as an endogenous variable for our outcomes of interest including regional income level, retail sales, employment rate, etc. One possible way to estimate CATE that can be used even in treatment effect models with endogenity is a marginal treatment effects (MTE) approach (Heckman and Vytlacil, 1999, 2005), and this is the approach adopted in this paper. Once MTE is estimated, it can be used to estimate CATE (and also many other treatment parameters including ATE and ATET). In addition, the estimated MTE can provide us with information on how the treatment effect varies across cities in terms of their unobserved characteristics. To our best knowledge, the unobserved heterogeneity in the effects of place-based policies is an issue not yet addressed in the literature.

For the conditioning variables in the CATE parameter, we focus on the following regional characteristics that can be classified into three categories: (1) demographic variables, (2) variables measuring the degree of suburbanization, and (3) local production-network variables. For the demographic variables, the population density and the share of the population older than 65 years are considered. For the suburbanization variables, we consider the number of cars per household and the proportion of large-scale retail stores over all retail stores. In addition, we consider the variables that represent the structure of local production-network to address the unique feature of the Japanese URZ program that the primary program benefits are limited to medium- and small-sized firms that belong to retail and service industries. That is, the beneficiaries of the primary benefits are limited in terms of both firm size and industry, let alone location. This would imply that the characteristics of the production networks to which those firms belong are important for region-wise policy impacts. We can construct these characteristic variables by using a unique inter-firm transaction data provided by a Japanese major credit research firm, Tokyo Shoko Research (TSR). In particular, we consider two kinds of measures to characterize the networks. First, we consider the ratio of the number of transactions within a city to the total number of transactions made by firms in the city, which we call the *degree of autarky*.

[3]Similarly, Lu et al. (2019) see how the treatment effects of China's place-based policy vary with the quality of infrastructure and the market potential, though they found no statistically significant differences.

[4]Besides these papers, most studies in the literature use more or less extended versions of the difference-in-difference method. As exceptions, Hanson (2009) estimates treatment effects of the federal urban empowerment zone program in USA with the instrumental variable method. He took Ways and Means membership and the number of years a member was on the committee at the time of designation as instruments. Also, some studies use the regression discontinuity design (RDD) (e.g., Freedman, 2015; Ehrlich and Seidel, 2018; Lu et al., 2019). Neumark and Young (2019) point out that we need to be cautious about the validity of the parallel trend assumption when using the DID method.

Second, we consider the industrial structure of city-level production networks by capturing the relative magnitude of within-industrial transactions and between-industrial transactions. We measure this by what is known as *modularity* in network science literature (Newman and Girvan, 2004).

To summarize, the objective of this paper is two-fold: (i) to examine the heterogeneity in the treatment effects of the Japanese URZ program using the MTE framework, and (ii) to propose empirically optimal treatment choice rules in terms of certain set of regional characteristic variables. For the outcome variables, we consider per-capita income to see the overall effects of the program and per-capita retail sales amount to see the effects on the main target industry of the program. Our main results are as follows. Whereas we consider three types of conditioning variables (i.e., demographic variables, suburbanization variables, and local production-network variables), our result indicates that the treatment choice based on the suburbanization variables is most successful for both outcomes. In particular, we find that cities where the number of cars per household is large and the fraction of large-scale stores is low should be targeted. For the program to be effective, the degree of suburbanization should not be so low because the marginal effects of revitalizing city centers is small otherwise. However, the chance of the program working may be small if large-scale retail stores are prevalent.

The rest of this paper is organized as follows. In Section 2, we describe backgrounds on the URZ program in Japan. Section 3 introduces our conditioning variables in the CATE parameter. In particular, we formally define the two measures to characterize the structure of local production networks, namely, the degree of autarky and modularity. Section 4 outlines our empirical strategy. In Section 5, we report our main empirical findings after describing the data used, and Section 6 concludes.

2 The urban revitalization zone program in Japan

The Act on Vitalization in City Center seeks to revitalize civic function and economic activities of city centers which have been deteriorated due to factors such as falling birthrate and aging population, motorization, and the expansion of large-scale retail stores[5]. The Act went into effect in 1998 under the necessity of protecting small-scale retail stores in city center after the government abolished the Large-scale Retail Stores Law, which controls the opening of large-scale retail stores, following the advice of the World Trade Organization (WTO). Under the Act, the central government supports local governments on the revitalization of city centers. A local government tries to achieve its goals through a variety of revitalization activities in a zone located in the city center, which we call urban revitalization zone (URZ). In the first several years, the program was operated under a notification system (i.e., local governments could receive program benefits as long as they submit notice). However, because the government found that the program had not been successful[6], it changed the operation of the program to an approval system in 2006.

Under the URZ program, the local governments can pursue the revitalization of their city centers

[5]This section is heavily based on 2018 report of the Board of Audit of Japan, https://www.jbaudit.go.jp/pr/kensa/result/30/pdf/301221_zenbun_01.pdf (in Japanese), accessed Octber 23, 2019.

[6]For details about the situation until 2006, see 2006 report of the Board of Audit of Japan, https://www.jbaudit.go.jp/report/zuiji/pdf/h18/request_h18.tmogai.pdf (in Japanese), accessed October 23, 2019.

through a wide variety of activities. According to the basic policy decided upon by the Cabinet[7], the program is organized into five categories: urban improvement, capacity improvement, housing, business assistance, and public infrastructure improvement. The urban improvement activities include the redevelopment of brownfields, the land readjustment for promoting the use of housing land, and the improvement of infrastructures such as streets, sewer systems, rivers, and parks. The capacity improvement activities include the construction of educational and cultural facilities, the construction of social welfare facilities, and encouraging citizen participation and community building. The housing activities include the development of public housing, the assistance for existing housing development, and the decrease in the rent of public housing. The public infrastructure improvement activities include the provision of the community bus service.

The business assistance activities are the most important here because they directly affect the economic impacts of the program. The activities include the following incentives. First, each URZ is eligible for funds that could be used for purposes such as consulting service, market research, and training program. Second, regulation for locations of large-scale retail stores in URZ is simplified. Third, subsidies are provided to nongovernmental projects such as the development of commercial facilities, which are typically carried out by associations composed of establishments in shop streets and private business operators such as urban development companies. Moreover, small and medium-sized firms in wholesale, retail, and service sectors that operate their business in such commercial facilities could benefit from loan with a preferential interest rate. Private business operators such as urban development companies that operate commercial facilities are also eligible for this benefit in addition to the subsidy. In a well-known case, an urban development company converted a vacant lot where a major private supermarket once stood to a facility that has a supermarket under its direct management and spaces for regional communities.

Applications for URZ designation are made at city level. The targeted zone is specified by the city government, which typically locates in a central business district of the city. The project period is also specified by the city government. It is typically (but not limited to) 5 years. It is possible that the city government applies for another project after it completed project(s). There is no clear-cut requirement to be eligible for URZ designation.[8] The Act states that there should be concerns of stagnation of economic activities in candidate zones in terms of land use and business activities. After the amendment of the Act in 2006, the city governments that apply for the program are required to provide supporting evidence for such concerns by using data of economic activities such as land use, the number of establishments, that of workers, that of retail stores, and retail sales. Typically, they show that values of those economic variables are declining over time. Moreover, they are required to set numerical targets. The URZ-designated city government has to report how much its goals have been achieved every year during its project period. The Act states that a project could be terminated

[7] https://www.kantei.go.jp/jp/singi/tiiki/chukatu/kettei/01_kihonhoushin190329.pdf (in Japanese), accessed October 30, 2019.

[8] This is in contrast to the New Markets Tax Credit program in USA studied by Freedman (2015), where qualification conditions include "tracts in metropolitan statistical areas (MSAs) with median family income (MFI) that does not exceed 80% of the greater of MSA MFI or statewide MFI qualify". On the other hand, this is similar to the Calfornia's enterprise zone program as Neumark and Kolko (2010) state "the process is not formulaic and appears to rely on subjective assessments".

depending on its progress even if its project period had not been completed. (See also Remark 4.1.) A city government, which has completed project(s), is not necessarily disqualified from applying for another project even if it has failed to achieve its goals, but the central government refers to the results of the completed project(s) when screening its application.

Table 1 shows the total numbers of URZs from 2007 to 2016 from our dataset, which is described in Section 5.1. In our dataset, the number of cities which has at least one approved URZ is 106 out of 661 cities. In other words, more than 15% of cities received the benefits from this policy.

Table 1: The total number of URZs

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
12	42	64	76	82	82	85	73	67	69

As is evident from the application process above, it is quite likely that the prospect of URZ designation is correlated with regional economic performance. We will address this issue with instrumental variable because we do not take the DID approach. Before the amendment of the Act, it was often the case that the city governments put out the preparation of application materials to private consulting firms. However, the amendment of the Act made the application process time-consuming and laborious. For example, it is reported that a city government spent much time on setting numerical targets. As we explain in Section 5, this motivates us to consider the average age of city government officers as the instrumental variable that directly affects the prospect of URZ designation but is only indirectly affected by regional economic outcomes.

Although the URZ-designated city governments have an obligation to achieve their numerical targets, it is within their discretion to choose economic variables on which they set numerical targets. As a result, project goals generally differs among URZs, but those that are commonly set are the numbers of workers and establishments (particularly for tertiary industries), the residential population, the retail sales, the shares of vacant lots and untenanted stores, the passenger volume of public transport such as bus and tram, the pedestrian traffic volume, and the number of visitors to public facilities such as station and tourist site.[9] In this paper, we consider the per-capita retail sale as an outcome variable because the retail sector is a major target of the program. Moreover, to see the general effects of the program, we also consider the per-capita income, although it is rarely set as a project goal.

We evaluate the policy with city-level variables, even though URZs are defined as areas which are strictly contained in cities. The primary reason is that the Act on Vitalization in City Center states that the program should be beneficial not only to targeted zones but also to their surrounding regions. Hence, it is not appropriate to focus solely on the targeted zones to evaluate the policy. Although the Act does not specify the surrounding regions, it is natural to consider a whole city because application to the program is made at city level. Moreover, we can obtain a sufficient number of observations even if we use city-level data because the program had a wide coverage (See Table 1). Hence, we abstract

[9]The city governments can also specify targeted zone. It is possible that one city has multiple zones, though it is quite rare.

away from the spillover effects within cities, which are major concerns in the literature (c.f., Neumark and Simpson, 2015).

It is important to see whether the total effect of the program, if any, outweighs its total cost. Table 2 summarizes information complied from the 2018 report of the Board of Audit of Japan on the total amount of money the URZ program had spent during our sample period. On average, the URZ program had spent around 310 billion yen, which is approximately worth 3.1 billion US dollars, per year.[10] Table 3 summarizes total expenditures over our sample period for each of the five categories[11]. The urban improvement activities naturally have the highest share because they involve the improvement of infrastructures which costs a lot of money.

Table 2: Total spending (million yen)

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
129,184	250,000	378,048	337,869	347,885	306,507	293,174	312,150	336,180	382,548

Table 3: Expenditure by category (million yen)

Category	Amount
Urban improvement	1,132,071 (41%)
Capacity improvement	639,419 (23.2%)
Housing	392,273 (14.2%)
Business assistance	388,613 (14.1%)
Public transportation improvement	203,393 (7.3%)

3 The choice of conditioning variables

In this paper, we classify regional characteristic variables into the following three categories: (1) demographic variables, (2) variables for the degree of suburbanization, and (3) local production-network variables. More detailed descriptions about these variables are as follows.

As the Act on Vitalization in City Center mentions a declining birthrate and aging population as a background for the necessity of revitalizing city centers, we consider the population density and the share of the population older than 65 years as demographic variables. The Act also mentions the change in consumer's lifestyle, which is often interpreted to indicate motorization and the expansion of large-scale retail stores. Hence, we consider the number of cars per household and the fraction

[10]The average yen-US dollar exchange rate over our sample period was 100.212 yen to the dollar. Hence, we assume that one US dollar is worth 100 yen when converting monetary units from yen to US dollar.

[11]The expenditures are calculated from 134 URZ plans which were assessed until 2016 in 90 cities, based on on-site investigation by the Board of Audit of Japan.

of large-scale retail stores over all retail stores as suburbanization variables. Here, large-scale retail stores are defined as stores with store floor area larger than 1,000 square meters.

In addition to the above, we consider the characteristics of local production networks. As we saw in Section 2, the centerpiece of the URZ program is the scheme that encourages the development of commercial facilities in targeted zones, where, in addition to the subsidies to local agents such as urban development companies that develop commercial facilities, firms that do their business in those facilities can receive financial assistance from Japan Finance Corporation through preferential interest rates. However, only small- and medium-sized firms of particular industries such as retail, wholesale, and service are eligible for the program benefits. This would imply that the characteristics of the production networks to which these firms belong are important for region-wise policy impacts.

This motivates us to include the characteristics of regional production networks in the conditioning variables. In particular, we consider two kinds of measures to characterize the networks. First, we consider the ratio of the number of transactions within a city to the total number of transactions made by firms in the city, which we call the *degree of autarky*. We consider this measure because only small firms are eligible for the program benefits. Even if small firms could get support from the program for starting new business, it would be important for their survival that productive firms locate nearby.[12] Because we do not have data on the productivities of individual firms, though, we presume that firms engaging in regional trade are more productive than those not doing so because engaging in regional trade entails sunk costs such as the cost of establishing extensive trade routes, and hence a city having a low degree of autarky would imply that many productive firms are there.[13] Moreover, because most of firms in commercial facilities are in retail and service sectors, it is important that there is enough local demand in the city. The productivity of local firms then matters because high productivity generally results in high wage that pushes up consumption for retail and service, which is related to *urbanization economies* due to Henderson (1986).[14]

Second, we consider the industrial structure of city-level production networks by looking at the relative magnitude of within-industrial transactions and between-industrial transactions. We measure this by what is known as *modularity* in network science literature (Newman and Girvan, 2004). The URZ program aims to make economic activities agglomerated in the central area rather than to make them dispersed all over the city. From perspective of economic theory, a rationale for this is the *agglomeration economies* that are positive externalities arising from economic agents locating close to each other. In the literature of urban and regional economics, the two types of agglomeration economies have been discussed: positive externalities caused by agglomeration of firms of the same industry, known as *localization economies*, and positive externalities caused by agglomeration of firms of different industries, known as *Jacobs externalities*.[15] Given that the program benefit is limited

[12] Givord et al. (2018) find that the French place-based policy, which provides incentives to small-sized firms, fails to keep the survival rate of small business low.

[13] The literature on international trade that looks at firm-level transactions (e.g., Bernard et al., 2007) shows that firms engaging in international trade are more productive that those not doing so. The sunk cost of engaging in international trade would be higher than that of engaging in regional trade due to cultural differences across countries.

[14] Another reason why the productivity of local firms is important is that local firms usually contribute money to local agents that operate commercial facilities. In general, subsidies are not enough for local agents to continue their projects. Hence, the existence of local firms that *can* contribute money is important for the sustainability of local projects.

[15] The localization economies are further divided into Marshall-Arrow-Romer (MAR) externality which emphasizes the

to particular industries, the relative magnitude of the two types of agglomeration economies is also important here, and it would be a critical factor for the effectiveness of the program. In fact, there is a large literature that estimates each type of agglomeration economies but results obtained so far are quite diverse as summarized in De Groot et al. (2016). This means that the relative magnitude of the two types of agglomeration economies is a regional characteristic that we should control.

In the subsequent subsection, we formally define the production network and the two network measures we consider to characterize the production networks.

3.1 Production networks

Let S be the set of firms. We define a production network by the adjacency matrix $A = [a_{ij}]_{i,j\in S}$ where

$$a_{ij} = \begin{cases} 1 & \text{if firm } j \text{ is a buyer of firm } i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

That is, if $a_{ij} = 1$, commodity goes from firm i to firm j whereas money goes from firm j to firm i. For $i \in S$, we define $d_{Ii} \equiv \sum_{j\in S} a_{ji}$ as the *in-degree* of firm i. This is the total number of firm i's sellers. Similarly, for $i \in S$, we define $d_{Oi} \equiv \sum_{j\in S} a_{ij}$ as the *out-degree* of firm i. This is the total number of firm i's buyers. We also call $d_i \equiv \frac{d_{Ii}+d_{Oi}}{2}$ the *degree* of firm i. Note that A is generally not symmetric. That is, even if firm j is a buyer of firm i, firm i is not necessarily a buyer of firm j. Hence, in general, $d_{Ii} \neq d_{Oi}$.

As discussed above, we consider the two regional characteristics of the production network, the degree of autarky and the modularity, each of which is now defined.

First, the *degree of autarky* of city k is defined below:

$$h^k = \frac{\sum_{i\in S^k}\sum_{j\in S^k} a_{ij}}{\frac{1}{2}\left(\sum_{i\in S^k}\sum_{j\in S} a_{ij} + \sum_{i\in S^k}\sum_{j\in S} a_{ji}\right)} = \frac{\sum_{i\in S^k}\sum_{j\in S^k} a_{ij}}{\sum_{i\in S^k} d_i}, \tag{3.2}$$

where $S^k \subseteq S$ is the set of firms in city k. The numerator is the number of within-city transactions whereas the denominator is the total number of transactions made by city k's firms which include transactions with other cities. This is exactly the *homophily index* of Currarini et al. (2009) who consider networks of friendship. We presume that a city with lower degree of autarky has higher regional productivity.

Second, to capture the relative importance of the two types of agglomeration economies (i.e., localization economies and Jacobs externalities), we divide firms in a city into several groups according to their industry so that $S^k_\ell \subseteq S^k$ is the set of firms of industry ℓ in city k. Because agglomeration economies are local externalities, we restrict our attention to industrial structures of local production networks. That is, for each city, we consider a production network limited to transactions within the city. Formally, we define the production network of city k by $A^k = [a_{ij}]_{i,j\in S^k}$. We define in-degrees, out-degrees, and degrees of firms in city k's local production network in an anlogue manner to the

role of knowledge spillover and Porter's externality which emphasizes the role of competition.

national network, i.e., $d^k_{Ii} = \sum_{j \in S^k} a_{ji}, d^k_{Oi} = \sum_{j \in S^k} a_{ij}$, and $d^k_i = \frac{d^k_{Ii} + d^k_{Oi}}{2}$ for $i \in S^k$. We then define the *modularity* of city k by

$$Q^k = \frac{1}{M^k} \sum_{\ell} \left(M^k_\ell - E[M^k_\ell | d^k] \right), \tag{3.3}$$

where $M^k_\ell = \sum_{i,j \in S^k_\ell} a_{ij}$ (i.e., the total number of transactions within industry ℓ of city k), $M^k = \sum_\ell M^k_\ell$ (i.e., the total number of transactions within city k), $d^k = \{d^k_{Ii}, d^k_{Oi}\}_{i \in S^k}$ is the degree distribution in city k, and $E[M^k_\ell | d^k]$ is the conditional expectation of M^k_ℓ when links are formed at random given degree distribution d^k. It follows that $E[M^k_\ell | d^k] = \frac{1}{M^k} \left(\sum_{i \in S^k_\ell} d^k_{Ii} \right) \left(\sum_{i \in S^k_\ell} d^k_{Oi} \right)$.

As is clear from the definition, the modularity is large [resp. small] when within-industry transaction is large [resp. small] compared to the case where transactions are randomly assigned given degree distribution. Therefore, the modularity captures the relative importance of localization economies and Jacobs externalities economies. In particular, the modularity is large [resp. small] when localization economies [resp. Jacobs externalities] are the stronger.

4 Estimation of conditional treatment effects

In this section, we describe our econometric approach for estimating possibly heterogeneous treatment effects of the URZ program. The parameter of primary interest to be estimated is the average treatment effects conditional on a set of characteristics variables of cities X, say $CATE(X)$. Since we should suspect the presence of endogeneity in the URZ program assignment process, using matching or propensity score methods is inappropriate. Then, to accomplish this task, this study adopts the MTE framework (Heckman and Vytlacil, 1999, 2005).

4.1 Marginal treatment effects

Suppose that we have panel data of N cities over T time periods: $\{(Y_{it}, D_{it}, X_{it}, Z_{it})\colon i = 1, \ldots, N, t = 1, \ldots, T\}$. Here, Y_{it} is the outcome variable of interest (e.g., par-capita income and per-capita retail sales), and D_{it} is the treatment variable which takes $D_{it} = 1$ if city i is in the URZ program in year t, $D_{it} = 0$ otherwise. X_{it} is the vector of covariates that affect Y_{it}, where $X_{it} = (X^{dem}_{it}, X^{net}_{it}, X^{sub}_{it})$ with X^{dem}_{it}, X^{net}_{it} and X^{sub}_{it} being the sets of demographic variables, production-network variables and suburbanization variables, respectively. For simplicity, we assume that the elements of X_{it} are exogenous. Z_{it} is the vector of instrumental variables that are direct determinants of D_{it}, but indirectly affect Y_{it} only through D_{it}.

Let $Y_{it}(d)$ denote the potential outcome when $D_{it} = d$ so that $Y_{it} = D_{it} Y_{it}(1) + (1 - D_{it}) Y_{it}(0)$. In order to capture the effects of X_{it} on $Y_{it}(d)$ in a flexible functional form, we assume the following generalized additive potential outcome model with the dynamic latent index formulation for the treatment assignment:

$$\begin{aligned} Y_{it}(d) &= g^{dem}_d(X^{dem}_{it}) + g^{net}_d(X^{net}_{it}) + g^{sub}_d(X^{sub}_{it}) + \varepsilon_{it}(d), \quad \text{for } d \in \{0, 1\} \\ D_{it} &= \mathbf{1} \left(\rho D_{i,t-1} + h^{dem}(X^{dem}_{it}) + h^{net}(X^{net}_{it}) + h^{sub}(X^{sub}_{it}) + Z^\top_{it} \gamma + \kappa_i \geq u_{it} \right), \end{aligned} \tag{4.1}$$

where $\varepsilon_{it}(d)$ is an unobservable determinant of $Y_{it}(d)$, u_{it} is an unobservable cost of taking $D_{it} = 1$, and κ_i is an individual-specific effect on D_{it}. Here, ρ, γ, g_0^{c}, g_1^{c} and h^{c}, where c = "*dem*", "*net*", "*sub*", are unknown parameters to be estimated. One can view the model in (4.1) as a dynamic-panel extension of the latent index model of Heckman and Vytlacil (1999, 2005). Note that the initial condition $D_{i0} = 0$ holds deterministically for all i's since the URZ program started at $t = 1$.

Remark 4.1. While our treatment assignment model assumes that URZ designation is judged every year t, in reality, the program is normally effective for five years once accepted. However, the law authorizing the program states that the national government requests the URZ-designated cities to submit the progress report every year. Then based on the report, the government may cancel the URZ designation when the effectiveness of the program is questionable. Thus, our model (4.1), which decides the treatment status D in every t, may be justifiable in this respect.

Here, we assume that u_{it}'s are independent and identically distributed as the standard normal $N(0, 1)$ both in time and across cities. In addition, we assume that while u_{it} is allowed to be correlated with $\varepsilon_{it}(d)$ (which is the source of endogeneity), $u_{it'}$'s are independent of $\varepsilon_{it}(d)$ if $t' \neq t$ (which ensures the independence between $D_{i,t-1}$ and $\varepsilon_{it}(d)$). Similarly, we assume that κ_i's are identically distributed as normal with mean 0 and standard deviation σ_κ, $N(0, \sigma_\kappa^2)$, independently of all independent variables and error terms. Then, due to the reproductive property of the normal distribution, $u_{it} - \kappa_i$ is distributed as $N(0, 1 + \sigma_\kappa^2)$. Thus, we can re-write the treatment assignment model as

$$D_{it} = \mathbf{1}\left(P_{it} \geq V_{it}\right),$$

where

$$P_{it} = \Phi\left(\frac{\rho D_{i,t-1} + h^{dem}(X_{it}^{dem}) + h^{net}(X_{it}^{net}) + h^{sub}(X_{it}^{sub}) + Z_{it}^\top \gamma}{\sqrt{1 + \sigma_\kappa^2}}\right),$$

$$V_{it} = \Phi\left(\frac{u_{it} - \kappa_i}{\sqrt{1 + \sigma_\kappa^2}}\right) \sim \text{Uniform}[0, 1],$$

and $\Phi(\cdot)$ denotes the standard normal distribution function.[16] Then, the MTE of the URZ program on Y_{it} is defined as

$$\begin{aligned} MTE(x^{dem}, x^{net}, x^{sub}, p) &\equiv E[Y_{it}(1) - Y_{it}(0) | X_{it}^{dem} = x^{dem}, X_{it}^{net} = x^{net}, X_{it}^{sub} = x^{sub}, V_{it} = p] \\ &= \sum_{\mathrm{c} \in \{\text{“}dem\text{”}, \text{“}net\text{”}, \text{“}sub\text{”}\}} [g_1^{\mathrm{c}}(x^{\mathrm{c}}) - g_0^{\mathrm{c}}(x^{\mathrm{c}})] + E[\varepsilon_{it}(1) - \varepsilon_{it}(0) | V_{it} = p] \end{aligned} \tag{4.2}$$

under the independence between $\varepsilon_{it}(d)$ and X_{it}. Recall that V_{it} can be interpreted as the implicit cost of taking the treatment. Hence, if p is small [resp. large], $MTE(x^{dem}, x^{net}, x^{sub}, p)$ is the expected treatment effect on cities with observable characteristics $(x^{dem}, x^{net}, x^{sub})$ that are likely [resp.

[16] It should be important to note that the assumptions introduced above ensures that P_{it} is independent of $\varepsilon_{it}(d)$. However, it is not independent of V_{it} due to the individual effect κ_i, except for the case of $t = 1$. Because of this dependence between P_{it} and V_{it}, semiparametric approaches proposed in the literature (e.g., Carneiro and Lee, 2009) cannot be directly applied here, and hence we use a parametric approach as described below.

unlikely] to be treated.

As mentioned in the introduction, once MTE is estimated, we can use it to calculate many other treatment effect parameters (see, e.g., Heckman and Vytlacil, 1999, 2005), including CATE. Since V_{it} is distributed as Uniform$[0,1]$, we can obtain $CATE(x^{dem}, x^{net}, x^{sub})$ in the following manner:

$$CATE(x^{dem}, x^{net}, x^{sub}) = \int_0^1 MTE(x^{dem}, x^{net}, x^{sub}, p)\mathrm{d}p.$$

From this, the estimation of ATE is rather straightforward: $ATE = E[CATE(X_{it}^{dem}, X_{it}^{net}, X_{it}^{sub})]$.

4.2 Maximum likelihood estimation

To estimate our model, we use a maximum likelihood approach. Suppose that $\varepsilon_{it}(d)$ and u_{it} are jointly normally distributed as

$$\begin{pmatrix} \varepsilon_{it}(d) \\ u_{it} \end{pmatrix} \sim N\left(\mathbf{0}, \Sigma_{du}\right), \quad \text{where} \quad \Sigma_{du} \equiv \begin{pmatrix} \sigma_d^2 & \sigma_{du} \\ \sigma_{du} & 1 \end{pmatrix}$$

for $d \in \{0,1\}$. In addition, we assume that the functions $g_d^{\mathrm{c}}(\cdot)$'s can be characterized by a fixed number of basis functions, such as power series, splines, Fourier series, etc. For example,

$$g_d^{\mathrm{c}}(x^{\mathrm{c}}) \equiv g_d^{\mathrm{c}}(x^{\mathrm{c}}; \delta_d^{\mathrm{c}}) = \sum_{j=1}^{J} b_j(x^{\mathrm{c}})\delta_{d,j}^{\mathrm{c}}$$

with $\delta_d^{\mathrm{c}} = (\delta_{d,1}^{\mathrm{c}}, \dots, \delta_{d,J}^{\mathrm{c}})^\top$. More specifically, for c = "*net*", letting $x^{net} = (x_a, x_m)^\top$, where x_a and x_m are the degree of autarky and the modularity of a production network as defined in Section 3.1, respectively, the second-order power-series formulation is given as $g_d^{net}(x^{net}) = x_a\delta_{d,1}^{net} + x_m\delta_{d,2}^{net} + x_a^2\delta_{d,3}^{net} + x_a x_m\delta_{d,4}^{net} + x_m^2\delta_{d,5}^{net}$. Similarly, we assume that $h^{\mathrm{c}}(\cdot)$'s can be characterized by a set of basis functions such that $h^{\mathrm{c}}(x^{\mathrm{c}}) \equiv h^{\mathrm{c}}(x^{\mathrm{c}}; \alpha^{\mathrm{c}})$ with α^{c} being the vector of corresponding basis coefficients.

Let the set of all unknown parameters to be estimated be Θ, which is comprised of $(\delta_0^{\mathrm{c}}, \delta_1^{\mathrm{c}}, \alpha^{\mathrm{c}})$ for c = "*dem*", "*net*", "*sub*", (ρ, γ) and $(\sigma_1, \sigma_0, \sigma_{1u}, \sigma_{0u}, \sigma_\kappa)$. Further, let $\phi(\cdot; \sigma^2)$ denote the normal density with mean zero and variance σ^2, and $\phi_2(\cdot, \cdot; \Sigma)$ be the bivariate normal density with mean zero and covariance matrix Σ. Then, the likelihood function for an observation (i,t) conditional on κ is given by

$$K_{it}(\Theta, \kappa) \equiv \left[\int^{S_{it}(\kappa)} \phi_2(e_{it}(1), s; \Sigma_{1u})\mathrm{d}s\right]^{D_{it}} \left[\int_{S_{it}(\kappa)} \phi_2(e_{it}(0), s; \Sigma_{0u})\mathrm{d}s\right]^{1-D_{it}},$$

where

$$e_{it}(d) \equiv Y_{it} - g_d^{dem}(X_{it}^{dem}; \delta_d^{dem}) - g_d^{net}(X_{it}^{net}; \delta_d^{net}) - g_d^{sub}(X_{it}^{sub}; \delta_d^{sub}) \text{ for } d \in \{0,1\}$$
$$S_{it}(\kappa) \equiv \rho D_{i,t-1} + h^{dem}(X_{it}^{dem}) + h^{net}(X_{it}^{net}) + h^{sub}(X_{it}^{sub}) + Z_{it}^\top\gamma + \kappa.$$

Hence, letting

$$\bar{K}_i(\Theta) \equiv \int \prod_{t=1}^{T} K_{it}(\Theta, \kappa)\phi(\kappa; \sigma_\kappa^2)\mathrm{d}\kappa, \tag{4.3}$$

the maximum likelihood estimator of Θ can be obtained by solving $\max_\Theta \sum_{i=1}^{N} \log \bar{K}_i(\Theta)$. However note that to evaluate the likelihood function given in (4.3), the integration with no closed-form solution must be solved. Thus, in application, we use a simulated maximum likelihood (SML) method to estimate Θ with a Monte Carlo approximation. Specifically, we replace $\bar{K}_i(\Theta)$ by $\hat{K}_i(\Theta)$, where

$$\hat{K}_i(\Theta) \equiv \frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T} K_{it}(\Theta, \kappa_r),$$

κ_r's are random numbers drawn from $N(0, \sigma_\kappa^2)$, and R is the total number of the drawn random numbers. Then, our SML estimator is defined as the maximizer of the simulated log-likelihood function $\sum_{i=1}^{N} \log \hat{K}_i(\Theta)$.

4.3 Estimation of MTE

In order to estimate the MTE parameter (4.2), we need to estimate the conditional expectation of $\varepsilon_{it}(d)$ given $V_{it} = p$ for both $d = 1$ and 0. Let $U_1 = \varepsilon_{it}(1)/\sigma_1$ and $U_2 = (u_{it} - \kappa_i)/\sqrt{1 + \sigma_\kappa^2}$. Then, both U_1 and U_2 are distributed as standard normal with correlation coefficient

$$Cor(U_1, U_2) = \frac{\sigma_{1u}}{\sigma_1\sqrt{1 + \sigma_\kappa^2}},$$

as we have assumed that κ_i is independent of $\varepsilon_{it}(1)$. Using the property of the standard bivariate normal distribution, we have

$$E(U_1|U_2 = u_2) = \frac{\sigma_{1u}}{\sigma_1\sqrt{1 + \sigma_\kappa^2}}u_2.$$

Then, noting the equality $E(\varepsilon_{it}(1)|V_{it} = p) = \sigma_1 E(U_1|U_2 = \Phi^{-1}(p))$, we have

$$E(\varepsilon_{it}(1)|V_{it} = p) = \frac{\sigma_{1u}}{\sqrt{1 + \sigma_\kappa^2}}\Phi^{-1}(p).$$

By the same argument, we also have $E(\varepsilon_{it}(0)|V_{it} = p) = \sigma_{0u}\Phi^{-1}(p)/\sqrt{1 + \sigma_\kappa^2}$. From these results, we can write

$$MTE(x^{dem}, x^{net}, x^{sub}, p) = \sum_{\mathrm{c} \in \{\text{“}dem\text{”}, \text{“}net\text{”}, \text{“}sub\text{”}\}} [g_1^{\mathrm{c}}(x^{\mathrm{c}}) - g_0^{\mathrm{c}}(x^{\mathrm{c}})] + \frac{\sigma_{1u} - \sigma_{0u}}{\sqrt{1 + \sigma_\kappa^2}}\Phi^{-1}(p).$$

Finally, our MTE estimator $\widehat{MTE}(x^{dem}, x^{net}, x^{sub}, p)$ of $MTE(x^{dem}, x^{net}, x^{sub}, p)$ can be obtained by replacing the unknown parameters by their SML estimators.

4.4 Optimal treatment assignment

As in Kitagawa and Tetenov (2018), an socially optimal treatment rule from the utilitarian perspective can be written as

$$G^* \in \arg\max_{G \in \mathcal{G}} E[CATE(X) \cdot \mathbf{1}\{X \in G\}],$$

where G is the so-called *decision set*, which is a subset of the support of X, and $\mathcal{G}$ is a collection of candidate decision sets. A sample analogue of the above maximization problem is called the Empirical Welfare Maximization (EWM). The EWM problem in our context would be given as follows:[17]

$$\hat{G}_{EWM} \in \arg\max_{G \in \mathcal{G}} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} [CATE(X_{it}) \cdot \mathbf{1}\{X_{it} \in G\}].$$

However, since the true $CATE(X_{it})$ is unknown, the above EWM rule $\hat{G}_{EWM}$ is not feasible. Then, we consider a feasible version of the EWM rule, in which $CATE(X_{it})$ is replaced by its estimate, which corresponds to the *m-hybrid* EWM approach of Kitagawa and Tetenov (2018):

$$\hat{G}_{m-hybrid} \in \arg\max_{G \in \mathcal{G}} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} [\widehat{CATE}(X_{it}) \cdot \mathbf{1}\{X_{it} \in G\}], \tag{4.4}$$

where

$$\widehat{CATE}(x) \equiv \int_0^1 \widehat{MTE}(x^{dem}, x^{net}, x^{sub}, p)\mathrm{d}p$$

with $x = (x^{dem}, x^{net}, x^{sub})$. Practical problems for the EWM approach are: when the set $\mathcal{G}$ is complicated, (1) solving the above maximization problem is computationally difficult, and (2) there is a difficulty in the interpretation of the resulting treatment rule. Thus, in order to facilitate the analysis, we consider the following quadrangle treatment rule:

$$\mathcal{G}^{\mathrm{c}}_{quad} \equiv \left\{ x : \beta^l_{\mathrm{c}} \le x^{\mathrm{c}} \le \beta^u_{\mathrm{c}}, \ \ (\beta^l_{\mathrm{c}}, \beta^u_{\mathrm{c}}) \in \mathbb{R}^{2\dim(x^{\mathrm{c}})} \right\} \tag{4.5}$$

for each c = "*dem*", "*net*", "*sub*". Then, we compute three different (sub)optimal treatment assignment rules based on $\mathcal{G}^{dem}_{quad}$, $\mathcal{G}^{net}_{quad}$ and $\mathcal{G}^{sub}_{quad}$, respectively, and discuss which type of characteristic variables would be the most effective to construct a socially optimal assignment rule for the URZ program in Japan.

Remark 4.2. As pointed out in Kitagawa and Tetenov (2017), the EWM-based (i.e., utilitarian) treatment assignment may result in an "unfair" treatment rule since only the gains from the treatment matter in specifying the treatment rule. Recalling that the first objective of the Japanese URZ program is to prevent the deterioration of local economies, the cities that should be treated in terms

[17] Note that the treatment rule $\hat{G}_{EWM}$ is a "pooled" rule pooled over the whole time period. If one is interested in an optimal treatment rule for each period, one may modify the definition as $\hat{G}_{t,EWM} \in \arg\max_{G \in \mathcal{G}} n^{-1} \sum_{i=1}^{n} [CATE(X_{it}) \cdot \mathbf{1}\{X_{it} \in G\}]$.

of the policy objective may not necessarily coincide with those selected in the "optimal" treatment group. To overcome this problem, Kitagawa and Tetenov (2017) proposes an *equally-minded* treatment assignment rule; however, we leave the application of this new approach to future work.

5 Empirical analysis

5.1 Data

5.1.1 Definition of sample and treatment

For our empirical analysis, we construct an annual balanced panel dataset of all Japanese cities between 2007 and 2016. We eliminate Tokyo Ward District and cities which experienced merger during the research period from our sample. Also, we drop cities with missing information[18].

Our treatment variable takes unity if a city is assigned as an URZ and takes zero otherwise. For the first year of the assignment, we set the variable takes one only if the assignment is done until September, because there might be some time lag until effects of the policy would be realized, if any. We already show the number of treated prefectures in Table 1 in Section 2.

5.1.2 Variable definition

For dependent variables Y, we adopt two variables as discussed in Section 2. The first variable is the per-capita income which shows the general effects of the policy. It is defined as the amount of taxable income over the number of individual tax payers.[19] The second variable is the per-capita retail sale which is directly related to the policy targeting. For standardization, the amounts of sales are divided by the population in the city.[20] Because we do not observe the variables for all years, we impute the values using weighted averages for the missing years.

Although it is not strictly required for identification under our parametric assumption, we include an instrumental variable which control the endogeneity between the treatment and the dependent variables. We consider the average age of city officials in general administrative positions as a factor which is related to regional politics and administration.[21] This factor affects the negotiation power of cities, which are likely to be correlated with the treatment as follows. In the assessment of the URZ

[18] Yubari, Akabira, and Utashinai cities have several variables which are not opened to public in 2016 Economic Census for Business Activity, because the number of offices are too small to ensure privacy. Rikuzentakata city has missing variables in many sources because of the East-Japan Earthquake.

[19] The information is taken from Annual Statistics on Municipal Tax by Japan Ministry of Internal Affairs and Communications.

[20] The population is taken from Vital Statistics, Ministry of Health, Labour and Welfare. The other variables are taken from the Census of Commerce, Ministry of Economy, Trade and Industry, for 2004, 2007, 2012 and 2014; the Economic Census for Business Frame, Ministry of Internal Affairs and Communications and Ministry of Economy, Trade and Industry, for 2009; and the Economic Census for Business Activity, Ministry of Internal Affairs and Communications and Ministry of Economy, Trade and Industry, for 2016. The amounts of sales for 2009 are missing. In our research period, the Japanese government was in process of replacing several economic surveys including the Census of Commerce by an comprehensive survey, the Economic Census for Business Frame and the Economic Census for Business Activity. Due to this transition, there are such complication about resource of our dataset.

[21] This information is taken from the Survey of Local Public Service Salary by the Ministry of Internal Affairs and Communications. We also have data on the number of city officials in general administrative positions, but it is highly correlated with population and thus contains little additional information.

program since 2006, the city governments are required to have prior negotiations with the national government. For this stage, cities need to prepare the basic plans by themselves.

This situation is different from that before 2006, when many cities could outsource the planning to external think tanks (Yokomori et al., 2008). Thus, city's ability of policy making might be an important factor to be treated. On the other hand, there is no clear path that relates these factors to regional economic status.

As discussed in Section 3, our explanatory variables consist of demographic variables X^{dem}, suburbanization variables X^{sub}, and network variables X^{net}. The demographic variables are population density, which is population divided by inhabitable area, and the fraction of population older than 65 years[22]. For the suburbanization variables, we incorporate the number of cars per household[23] and the fraction of large-scale retail stores over all retail stores[24].

For the network variables, we implement the degree of autarky and the modularity, as defined in Section 3.1. These variables are calculated from the firm-level micro data provided by a Japanese major credit research firm, Tokyo Shoko Research (TSR). We have data for 2006, 2011, 2012, and 2014. Thus we impute the values using weighted averages of the missing years. The data we use covers about 15% of all firms in Japan. TSR asks firms their important buyers and sellers. The firms rank their buyers and sellers up to 24th respectively. When constructing a production network as in (3.1), we let $a_{ij} = 1$ if firm i lists firm j as its important buyer, even if firm j does not list firm i as its important seller, and vice versa. The data also contains information such as address, up to three industries (4-digit), sales, and the number of employees for each firm. We use the information of address for calculating the degree of autarky and that of address and industry for calculating the modularity for each city. Table 4 summarizes the number of firms and the average degree, which is defined in Section 3.1, for each year, where firms that do not trade with others at all are excluded.

	2006	2011	2012	2014
# firms	440,521	636,344	603,014	678,239
Average degree	3.2334	3.286	3.4787	3.2934

Table 4: The number of firms and the average degree of firms in TSR data where firms that do not trade with others at all are excluded.

[22]Both variables are taken from the Vital Statistics. We took the natural logarithm of the population density.

[23]This information is taken from the Monthly Report on Car Ownership by The Automobile Inspection and Registration Information Association

[24]This information is taken from the Survey of Large-Scale Retail Stores by Toyo Keizai Shimposha.

	All		Control		Treated	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Treatment	0.10	(0.30)				
Instrument						
Age of Officers	43.02	(1.78)	43.06	(1.80)	42.67	(1.53)
Demographic						
Population density	20.51	(24.88)	20.90	(25.49)	16.98	(18.11)
Elderly population	0.26	(0.05)	0.26	(0.05)	0.26	(0.04)
Suburbanization						
Cars	0.33	(0.06)	0.33	(0.06)	0.34	(0.06)
Large retail stores	0.02	(0.01)	0.02	(0.01)	0.03	(0.01)
Network						
Autarky	0.37	(0.16)	0.36	(0.16)	0.48	(0.13)
Modularity	0.07	(0.05)	0.07	(0.05)	0.08	(0.04)
Y						
Income pc	2.95	(0.45)	2.95	(0.46)	2.92	(0.35)
Retail sales pc	0.88	(0.22)	0.86	(0.22)	0.99	(0.19)
Observations	6,610		5,958		652	

pc: per capita

Table 5: Descriptive statistics. "Population density" is the population divided by inhabitable area. "Elderly population" is the fraction of population older than 65 years. "Cars" is the number of cars per household. "Large retail stores" is the fraction of retail stores with store floor area larger than 1,000 square meters over all retail stores.

Table 5 reports the descriptive statistics for our dataset. Our sample consists of 6,610 observations, i.e. 661 cities for 10 years. Among 6,610 observations, 652 observations are treated. The number of cities which has been at least once treated during the research period is 107.

Whereas there is no appreciable difference between the fractions of elderly persons, the population density is lower in the treatment group than in the control group, which is consistent with the tendency that large cities are hardly URZ designated. Regarding the suburbanization variables, both the number of cars per household and the fraction of large-scale retail stores are higher in the treatment group than in the control group, which means that URZ designations have been approved for cities where the degree of suburbanization is relatively high. Regarding the network variables, the degree of autarky is higher in the treatment group than in the control group. Hence, URZ cities are more economically isolated than non-URZ cities. Another implication is that, under our presumption discussed in Section 3, the city-level productivity is lower in URZ cities than in non-URZ cities, which is consistent with the program's intention of targeting economically under-performing places. The moduarity is higher in the treatment group than in the control group, which means that URZ cities are more industrially specialized in terms of inter-firm transactions.

5.2 Estimation results

We estimate our model (4.1) by the simulated maximum likelihood procedure explained in Section 4.2. Table 6 reports results for the lagged treatment and instrumental variables. The estimated effects of the both variables on the prospect of URZ designation are postive and significant at the 1% level, respectively. The positive effect of the lagged variable is as expected because, once an URZ designation is approved, it is effective for several years (typically five years). The positive effect of the instrumental variable, the average age of city officials, is also as expected because city officials are, on average, more experienced when their average age is higher, and the application process involves no light tasks such as setting numerical targets and having face-to-face meetings with central government officials as we discusssed in Sections 2 and 5.1.2.

Dependent variable: Treatment		Coef.	S.E.
	Constant	-9.826***	(0.179)
	ρ	1.349***	(0.193)
Instrument	Age of Officers	0.0105***	(0.0036)
Observations		6,610	

Table 6: Results for random effect probit. Year dummies are also included in estimation but not reported. Standard errors in parentheses, *** p < 0.01, ** p < 0.05, * p < 0.1.

5.2.1 Treatment effects

Figure 1 reports the marginal treatment effect for each outcome variable, where the shaded area represents its 95% confidence interval.[25]

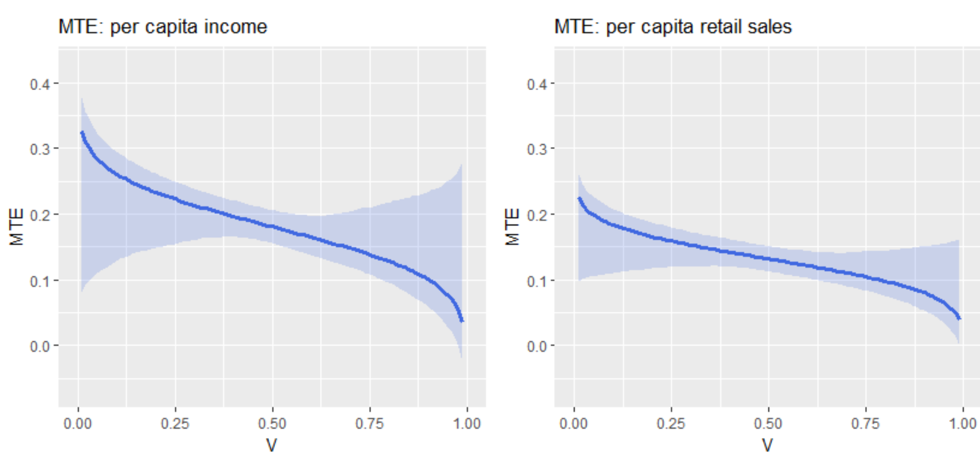

Figure 1: The marginal treatment effects (MTEs)

[25] The confidence intervals were calculated by using a parametric bootstrap with 3000 replications. In the estimation of the covariance σ_{du}, we parametrized it as $\sigma_{du} = \sigma_d \rho_{du}$ and estimated σ_d and ρ_{du} separately, where ρ_{du} is the correlation coefficient between $\varepsilon_{it}(d)$ and u_{it}. For computational simplicity, we further re-parametrised as $\rho_{du} = 2Logit(a_{du}) - 1$, and estimated a_{du} instead of directly estimating ρ_{du}, where $Logit(\cdot)$ is the standard logistic function, so that any candidate value of ρ_{du} is automatically included in the interval $(-1, 1)$. This is reflected in the asymmetry of the computed confidence intervals.

The x-axis is the unobservable resistance to the treatment (i.e., V_{it}). In our context, V_{it} would reflect, for examples, factors such as a sense of involvement of local retailers and landowners and how much residents feel attachment for their city. In fact, it is reported that some cities face complaints from residents about the preferential treatment to city centers. Moreover, the ability of town managers, which is related to the factors that Kolko and Neumark (2010) controls, could also be included in V_{it}. In fact, many local governments have a hard time getting hold of good personnel who possess expertise in town management.[26] In any case, it is reasonable that estimated MTE courves are downward sloping. Recalling that, if treatment effects were homogeneous, MTE curves would be flat, Figure 1 shows evidence that (unobserved) heterogeneity of treatment effects indeed exist.

The estimated ATE for each outcome variable is reported in Table 7. The ATEs for both outcomes are positively significant: on average, URZ designation boosts per capita income by 0.156 million yen per year. This is approximately worth 1,560 US dollars. Similarly, URZ designation increases per capita retail sale by 0.117 million yen per year, which is approximately worth 1,170 US dollars.

Outcome variable	ATE	95% CI
Income pc	0.156	[0.128, 0.183]
Retail sales pc	0.117	[0.101, 0.134]

Table 7: Estimated ATEs

5.2.2 Quadrangle treatment rules

Now, we move on to the investigation of practically effective URZ assignment rules. Note that since we have estimated the CATE for almost of all cities to be positive, without any budget constraints, the best policy is simply to designate all of such cities as URZs. However, such policy is clearly unrealistic. Thus, we consider a hypothetical scenario that the number of URZs that the government could have designated was 650 during 2007–2016 (recall that the number of treated observations in our panel data is 652).

For the treatment assignment rules, we use the quadrangle treatment rules which are defined in (4.5). We compute the (sub)optimal treatment rules from three different perspectives, the city's demographic characteristics, the degree of suburbanization, and the network structure of local firms. The results are given in Figures 2 and 3. In the figures, the resulting total empirical welfare (i.e., $\sum_{i=1}^{n}\sum_{t=1}^{T}[\widehat{CATE}(X_{it}) \cdot \mathbf{1}\{X_{it} \in G\}]$) are reported as EW. For comparison, the values of EW of the current policy (i.e., $\sum_{i=1}^{n}\sum_{t=1}^{T}[\widehat{CATE}(X_{it}) \cdot D_{it}]$) are 81.593 and 78.050 when the outcome variables are Income pc and Retail sales pc, respectively. Then, we can find that all the URZ designation rules proposed here are better than the current policy. In particular, those based on the suburbanization variables are the most successful in terms of the empirical welfare, for both outcomes. Moreover, in either case, the empirical welfare is maximized when the number of cars per household is larger and the fraction of large retail stores is lower.

[26] According to the on-site investigation by the Board of Audit of Japan, 17 of 56 cities that do not hire town managers mention the difficulty of finding well-qualified persons as a reason of it.

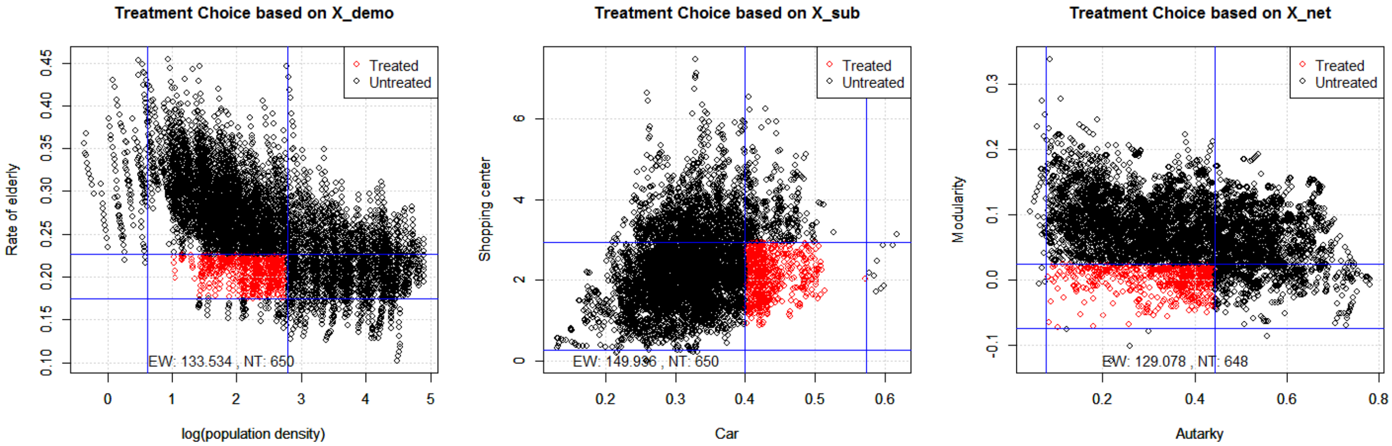

Figure 2: Quadrangle treatment rules (outcome: per-capita income)

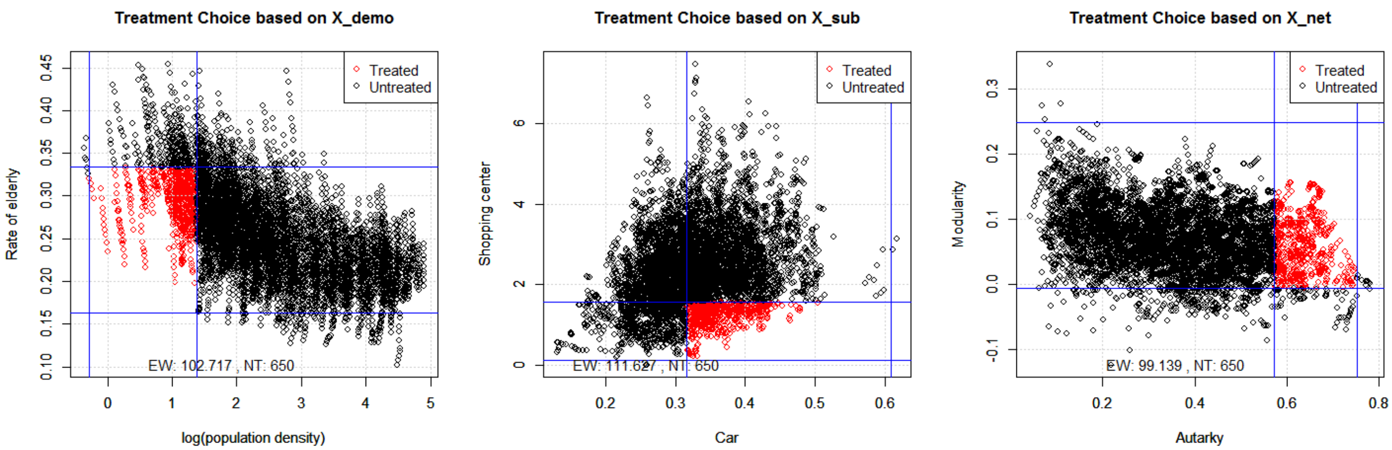

Figure 3: Quadrangle treatment rules (outcome: per-capita retail sales)

It would be quite reasonable that the program is more effective in cities where the fraction of large-scale retail stores is lower because it is difficult for newly developed commercial facilities in URZs to take away customers from preexisting large-scale retail stores especially when most households possess car(s).[27] Ironically, it is pointed out that the Act itself was also responsible for the expansion of large-scale retail stores. In the early stage of the program, the regulation of large-scale retail stores virtually did not exist, which resulted in their expansion into suburb areas. However, our finding implies that this conflicted with the goal of revitalizing city centers.[28]

[27] We need to be cautious when interpreting the treatment choice rules because the values of variables other than the treatment criterion variables X^{c} are not fixed in Figures 2 and 3. However, we obtained similar distributional patterns of $CATE(X^{\mathrm{c}})$ as in these figures, where the impacts of the variables excluding X^{c} are integrated out (see Figures 4 and 5 in Appendix A.1). Namely, the cities that should be treated in terms of X^{c} mostly overlap with those with higher $CATE(X^{\mathrm{c}})$.

[28] The government now promotes large-scale retail stores to locate in city centers aiming at the synergy effects after the amendment of the Act in 2006. It is worth investigating whether such synergy effects indeed exist. In fact, there is a case that developers such as railways companies develop shopping malls that are connected directly with stations and,

Our result also indicates that the number of cars per household should be large. In general, residents are more spatially dispersed in cities where the number of cars per household is larger. If both the number of cars per household and the fraction of large-retail stores are low, cities are already somehow compact. In such cities, the marginal effects of revitalizing city centers would not be so large. Hence, it is advised that policy makers target cities where residents are spatially dispersed but large-scale retail stores are not so prevalent. That is, targeting cities with the very high degree of suburbanization is not promising but targeting the ones with the very low degree of suburbanization is not desirable either to achieve enough effects.

A similar argument as above applies to the results regarding the demographic variables. For both outcome variables, the population density should not be high because the marginal benefits of revitalizing the city center are low otherwise. But still, cities need to have some potential for exploiting the program incentives, which leads to our finding that the share of elderly persons should not be high.

Regarding the network variables, we find that the degree of autarky should be low for per-capita income whereas it should be high for per-capita retail sale. For the per-capita income, the modurality, or the (inverse of) industrial diversity of inter-firm transactions, should be low, which means that Jacobs externalities are stronger than localization economies. This implies that positive shocks to retail and service sectors made by the program should spillover to other industries to boost the per-capita income. However, for such spillovers to work, the productivies of each firm need to be high, or the degree of autarky is low under our presumption, because to obtain spillover effects from different industries is more difficult than from the same industry. Our finding that Jacobs externality is the more important agglomeration economy is aligned with De Groot et al. (2016)'s conclusion "[o]verall, the support for Jacobs' hypothesis is relatively most convincing". For boosting the per-capita retail sales, it is important that local demands do not leak out, which would be the case when the city is isolated from others, or the degree of autarky is high. Our result indicates that this factor is more important than urbanization economies that we discussed in Section 3. On the other hand, it appears that the modurality is not really relevant for the per-capita retail sale.

6 Conclusion

In this paper, we estimated the treatment choice rules of the Japanese URZ program by extending the marginal treatment effect framework of Heckman and Vytlacil (1999, 2005) to a dynamic panel model to address the regional heterogeneity of the treatment effects of the program. According to our results, the government is likely to be successful in pushing up per-capita income or per-capita retail sales if it targets cities where the number of cars per household is high while the fraction of large-scale retail stores is low. We believe that our approach is capable of providing useful guidelines to policy administrators who decide which cities to target.

However, there are some limitation on our econometric approach that are left for future study. First, we did not considered dynamic models for the outcome equations while the treatment selection was modeled as a dynamic probit model. How to account for the dynamic aspects of potential

as a result, people do not go to the areas of shopping streets.

outcomes within an endogenous treatment model is still an open question with a few recent studies (e.g., Heckman et al., 2016, Han, Forthcoming). Second, although we have treated all elements of the covariates as exogenous variables, endogeneity of the network variables, for example, might be a concern. To address this issue we need additional instrumental variables, and the MTE model used should be modified accordingly. Third, the proposed treatment assignment rules considered only efficiency to improve overall social productivity. However, as mentioned in Section 2, one of the main goals of the URZ program is to revitalize the economy of inner-city commercial districts. In this respect, our treatment choice criterion is not necessarily optimal, and we should think of a treatment choice that prioritizes underperforming cities (see also the discussion in Remark 4.2). Finally, our econometric model ignored potential spatial externalities caused by the URZ program. It is reasonable to imagine that when an URZ program is introduced in a given city, then the surrounding cities not designated as URZs would also be affected. For instance, the local effects of place-based policy might be largely due to the local shift of economic activities rather than new economic activities. Estimation of effects of placed-based policies while taking such spatial externalities into account is an important recent research topic (e.g., Alonso et al., 2019, Ehrlich and Seidel, 2018, Givord et al., 2018, and Lu et al., 2019).

References

Alonso, J. M., R. Andrews, and V. Jorda (2019): "Do neighbourhood renewal programs reduce crime rates? Evidence from England," *Journal of Urban Economics*, 110, 51–69.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in international trade," *Journal of Economic Perspectives*, 21, 105–130.

Bhattacharya, D. and P. Dupas (2012): "Inferring welfare maximizing treatment assignment under budget constraints," *Journal of Econometrics*, 167, 168–196.

Briant, A., M. Lafourcade, and B. Schmutz (2015): "Can tax breaks beat geography? Lessons from the French enterprise zone experience," *American Economic Journal: Economic Policy*, 7, 88–124.

Carneiro, P. and S. Lee (2009): "Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality," *Journal of Econometrics*, 149, 191–208.

Chaurey, R. (2016): "Location-based tax incentives: Evidence from India," *Journal of Public Economics*.

Currarini, S., M. O. Jackson, and P. Pin (2009): "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica*, 77, 1003–1045.

De Groot, H. L., J. Poot, and M. J. Smit (2016): "Which agglomeration externalities matter most and why?" *Journal of Economic Surveys*, 30, 756–782.

Ehrlich, M. v. and T. Seidel (2018): "The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet," *American Economic Journal: Economic Policy*, 10, 344–74.

Freedman, M. (2015): "Place-based programs and the geographic dispersion of employment," *Regional Science and Urban Economics*, 53, 1–19.

Givord, P., S. Quantin, and C. Trevien (2018): "A long-term evaluation of the first generation of French urban enterprise zones," *Journal of Urban Economics*, 105, 149–161.

Han, S. (Forthcoming): "Identification in nonparametric models for dynamic treatment effects," *Journal of Econometrics*.

Hanson, A. (2009): "Local employment, poverty, and property value effects of geographically-targeted tax incentives: an instrumental variables approach," *Regional Science and Urban Economics*, 39, 721–731.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016): "Dynamic treatment effects," *Journal of Econometrics*, 191, 276–292.

Heckman, J. J. and E. Vytlacil (2005): "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 73, 669–738.

Heckman, J. J. and E. J. Vytlacil (1999): "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences*, 96, 4730–4734.

Henderson, J. V. (1986): "Efficiency of resource usage and city size," *Journal of Urban economics*, 19, 47–70.

Kitagawa, T. and A. Tetenov (2017): "Equality-minded treatment choice," *Cemmap Working Paper*.

——— (2018): "Who should be treated? empirical welfare maximization methods for treatment choice," *Econometrica*, 86, 591–616.

Kline, P. and E. Moretti (2014): "People, place, and public policy: some simple welfare economics of local economic development programs," *Annual Review of Economics*, 6, 629–662.

Kolko, J. and D. Neumark (2010): "Do some enterprise zones create jobs?" *Journal of Policy Analysis and management*, 29, 5–38.

Lu, Y., J. Wang, and L. Zhu (2019): "Place-Based Policies, Creation, and Agglomeration Economies: Evidence from China's Economic Zone Program," *American Economic Journal: Economic Policy*, 11, 325–60.

Manski, C. F. (2004): "Statistical treatment rules for heterogeneous populations," *Econometrica*, 72, 1221–1246.

Neumark, D. and J. Kolko (2010): "Do enterprise zones create jobs? Evidence from California' s enterprise zone program," *Journal of Urban Economics*, 68, 1–19.

Neumark, D. and H. Simpson (2015): "Place-based policies," in *Handbook of regional and urban economics*, Elsevier, vol. 5, 1197–1287.

Neumark, D. and T. Young (2019): "Enterprise zones, poverty, and labor market outcomes: Resolving conflicting evidence," *Regional Science and Urban Economics*, 78, 103462.

Newman, M. E. and M. Girvan (2004): "Finding and evaluating community structure in networks," *Physical review E*, 69, 026113.

Yokomori, T., K. Kuba, and Y. Nagasaka (2008): *Learning place-based policies from failed cases (in Japanese)*, Gakugei Shuppan.

Appendix

A.1 CATEs

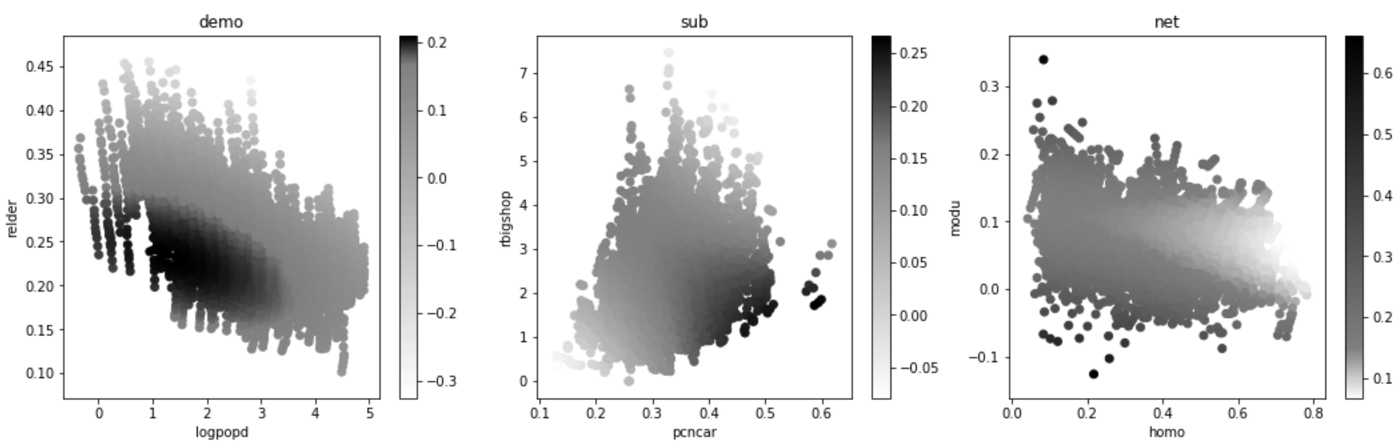

Figure 4: CATEs (outcome: per-capita income)

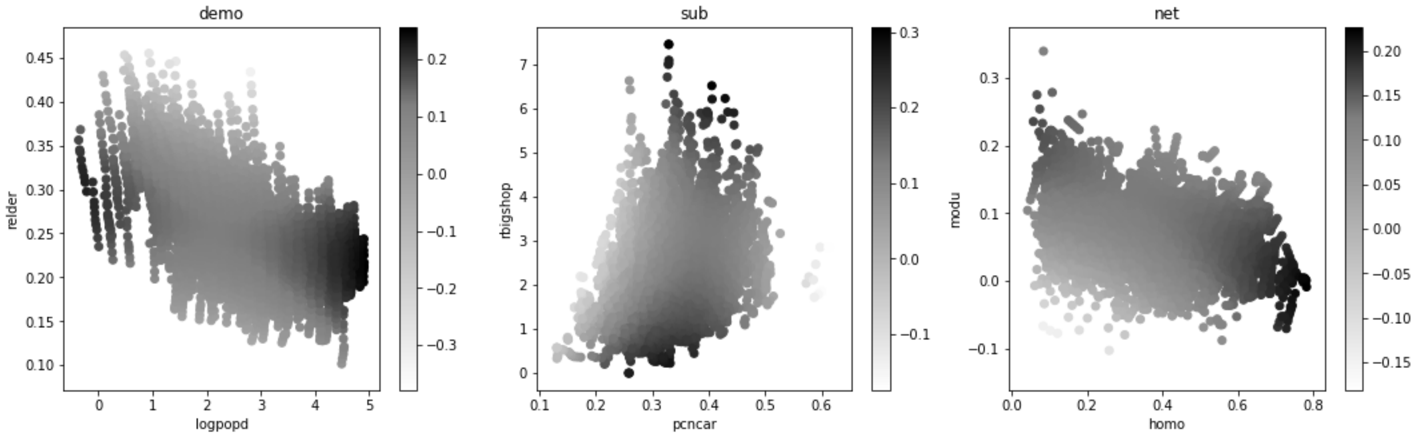

Figure 5: CATEs (outcome: per-capita retail sale)