博士論文

論文題目　Essays in Empirical Urban Economics

(実証都市経済学に関する研究)

氏名　岡本　千草

Contents

Introduction

People and economic activities are not equally distributed with respect to space. People often agglomerate in cities and this tendency is intensified in the global setting. The percentage of the world's population living in urban areas has been steadily increasing from 29.6% in 1950 to 53.9% in 2015 (United Nations, Department of Economic and Social Affairs, 2018). This figure is projected to reach 68.4% in 2050, indicating that more than two-thirds of the world's population will be living in urban areas within the next few decades. The number of megacities, that is, cities with more than 10 million inhabitants, has also doubled over the decades. Alongside this uneven distribution among cities, people and economic activities are not uniformly distributed even within cities. Cities often consist of specialized areas such as business, commercial, or residential districts. In residential districts, specific types of inhabitants often choose to locate within a certain area depending on their income level or race.

Urbanization has been promoting economic growth through an increase in productivity and innovation. However, over-agglomeration causes congestion, poor living environments, and inadequate infrastructure. The concentration of specific types of people in cities can be beneficial because they can gain access to thier networks (Bayer et al., 2008) and avoid negative neighborhood externalities. However, this unequal distribution also carries the risk of lousy segregation, which result in unequal job opportunities and amenities (Cutler et al., 2008).

What determines the spatial distribution of people and economic activities? To achieve optimal and sustainable spatial distribution via policy implementation, it is important to understand the determinants of existing distribution structures. This dissertation consists of three chapters. Each chapter individually focuses on one of three mechanisms that determine spatial distribution, as discussed in the literature: (1) change in production technology, (2) change in transportation costs, and (3) neighborhood externalities. Chapters 1 and 2 examine whether and how changes in production technology and transportation costs, respectively, alter the geographic distribution. Chapter 3 examines whether the externality effect truly exists.

Chapter 1 focuses on the change in production technology. Technological change, which is often represented as a shift in the production function, is one of the leading forces of economic growth. Many studies have addressed factor-neutral technological change—represented by the change in the total factor productivity (Solow, 1957)—which increases productivity but does not alter the proportion of production factors. In the last few decades, increasing wage inequality among skills has drawn attention to factor-biased technological change, such as skill-biased technological change (SBTC) (Violante, 2008). Under this change, some production factors (for example, skilled workers in SBTC) benefit more, but others (for example, unskilled workers in SBTC) do not. While technological change is identical in all regions, the impacts of the change can differ among regions depending on the regional factor endowment. Technological change will alter the spatial distribution of economic activities, which has been confirmed theoretically (Autor and Dorn, 2013; Accetturo et al., 2014; Tabuchi et al., 2018). Empirical studies also show that cities' or countries' skill endowments shape their economic growth or total factor productivity growth (Berman, 2000; Berger and Frey, 2016). This chapter focuses on a recent technological change—job process automation. It mainly affects labor productivity and can be classified as SBTC. By focusing on workers' migration decisions, this chapter examines the direct effect of the technological change on spatial distribution.

Chapter 2 conducts empirical analyses to examine the effect of new transportation infrastructures, especially the high-speed rail in Kyushu, Japan. This chapter focuses on the change in transportation costs. Unless the economy is self-sufficient, every area is connected with other areas. Firms transport inputs and outputs of production, and people themselves visit markets in other areas. Transportation infrastructures that reduce freight and passenger transportation costs are perceived as one of the main drivers of economic growth. Previous literature has mainly compared the effects of new transportation infrastructures on areas with and without the infrastructures, and many of these studies have found positive impacts on the local economy.[1] As this chapter discusses, transportation even reshapes the spatial distribution within connected areas. This phenomenon is well analyzed theoretically in New Economic Geography, pioneered by Krugman (1991), in which transportation costs are vital to determine the spatial distribution of activities. It shows that people agglomerate into one city when the transportation costs sufficiently decline. Although there are many empirical studies on transportation, this effect among connected areas has not been intensively studied with data.

The previous two chapters mainly focus on determinants of the spatial distribution of economic activities across cities. Chapter 3 focuses on neighborhood externalities as the main determinant of the distribution within a city. "Externalities arise when an agent does not compensate others for the effect of his actions (Kanemoto, 1980)." In a city where people agglomerate, there are a variety of externalities such as knowledge spillovers, traffic congestion, and pollution. This chapter primarily focuses on neighborhood externalities, which are externalities among different groups of people. The most typical ones are externalities between the rich and the poor and externalities among different races. Several studies have reported that neighborhood externalities determine the spatial distribution of different types of people (Kanemoto, 1980; Fujita, 1989). This chapter addresses a new type of neighborhood externalities in the literature, the "neighborhood externalities of one-room residents," and estimates whether the externalities actually exist. It enables us to discuss whether the externalities could be one of the determinants of the spatial distribution of people in a city.

The following provides the detailed summaries of each chapter.

Chapter 1: The effect of automation levels on US interstate migration[2]

This chapter investigates the extent to which job process automation, which has resulted in wage inequality and job polarization in the United States, has affected interstate migration over the past two decades. The level of automation in each state is calculated using data on the degree of automation of each occupation. In particular, this study examines how the difference in the levels among states explains the movement of migrants. The results show that people move to states with more automation in skilled occupations and less automation in unskilled occupations. This finding implies that automation has a complementary (substitution) effect

[1] Section 2.2 summarizes the literature.

[2] This chapter is based on following paper: Okamoto, Chigusa (2019) "The effect of automation levels on US interstate migration," *The Annals of Regional Science*, Vol.63, Issue 3, pp.519–539.

on skilled (unskilled) occupations. The results also show that the former effect is larger and more robust than the latter one. Further analyses use migration flow data classified into several subgroups and find that both skilled and unskilled workers in most occupations move to states with more automation in skilled occupations and less automation in unskilled occupations.

Chapter 2: Impacts of high-speed rail constructions on urban agglomeration: Evidence from Kyushu in Japan[3]

High-speed rail integrates urban and regional economies, and thus can possibly have significant impacts on the distribution of economic activities. Using the opening and extensions of a high-speed rail, Shinkansen, in Kyushu, Japan, we examine its effects on the distribution of economic activities across urban agglomerations. We focus on changes in land prices and estimate hedonic price equations to conduct a difference-in-differences analysis. We find that the large metropolitan areas gained from the high-speed rail by experiencing increases in land prices, whereas small metropolitan areas located between them lost by experiencing decreased land prices. However, such positive effects are shown to be limited to areas close to Shinkansen stations.

Chapter 3: Neighborhood externalities from "one-room apartments"

This study focuses on a new type of neighborhood externalities discussed in the literature–those from "one-room apartments," a term commonly used in Japan to refer to studio apartments. There is growing concern over the externalities caused by their occupants, and several municipalities in Japan have regulated the construction of new one-room apartments. This study examines whether negative externalities of one-room units toward non-one-room units exist by estimating the hedonic price equations with a rich dataset of rentals in Tokyo's 23 wards. We find clear evidence of negative neighborhood externalities both within apartments and within districts. This study addresses the externalities within a building, which are the most micro-level neighborhood externalities in the literature.

Acknowledgement

I would like to express my gratitude to my supervisor, Takatoshi Tabuchi, for his patient guidance and insightful comments. I am grateful to Yasuhiro Sato for his thoughtful comments and advice and for allowing me to use our joint work as a part of this dissertation. I would like to thank the other three dissertation committee members, Takaaki Takahashi, Kentaro Nakajima, and Mariko Nakagawa. I would also like to thank Atsushi Yamagishi, Chihiro Shimizu, Daiji Kawaguchi, Ikuto Aiba, Jiro Yoshida, Joshua Angrist, Kayo Tajima, Keisuke Kawata, Koichi Fukumura, Marcus Berliant, Masayuki Nakagawa, Michael Pflüger, Mitsuru Ota, Tomoya Mori and Yukiko Saito for their thoughtful comments (in alphabetical order). I gratefully acknowledge comments from participants of various seminars, workshops, and conferences.

[3]This chapter is a joint work with Yasuhiro Sato. This chapter is a revised version of following paper: Okamoto, Chigusa and Yasuhiro Sato (2018) "Impacts of high-speed rail construction on urban agglomerations: Evidence from Kyushu in Japan," CIRJE Discussion Paper Series, CIRJE-F-1097, University of Tokyo.

Chapter 2 was supported by JSPS KAKENHI Grant Numbers 15H03344, 17H02519, and 18H00842. Chapter 3 was supported by Joint Research Program No.821 at CSIS, UTokyo (Real Estate Database 1999-2016 by At Home Co.LTD).

Chapter 1

The effect of automation levels on US interstate migration

1.1 Introduction

Will computers ever replace humans in the workplace? The recent rapid progress in artificial intelligence (AI) has begun to suggest that production automation by AI and computers will proceed to the point that many jobs may be in danger of disappearing. According to Frey and Osborne (2017), about 47% of jobs are at risk of computerization. Widespread automation will change the content of people's jobs considerably, leading to alterations in both the labor market and society as a whole that "have the potential to disrupt the current livelihoods of millions of Americans" (Executive Office of the President of the United States of America, 2016). In particular, automation has resulted in wage inequality and job polarization, including wage and employment rises in the top and bottom of the income and skill distribution and a decline in the middle of the distributions (Autor, 2015).

While a number of studies have explored these phenomena theoretically and empirically in several countries (Autor et al., 2003; Goos and Manning, 2007; Autor and Dorn, 2013; Michaels et al., 2014; Autor, 2015), focusing on the extent to which automation affects national labor markets, the degree to which automation changes the population composition of local labor markets has been underexamined. For instance, Autor and Dorn (2013) and Accetturo et al. (2014) show that job polarization will occur in local labor markets that introduce automation widely. Autor and Dorn (2013) further test the implications from their model empirically. The recent study conducted by Acemoglu and Restrepo (2017) shows that exposure to robots negatively affects employment and wages across local labor markets in the United States and that the effects vary by such factors as education level. However, although they find that automation changes the occupational composition of local labor markets, the migration decision in response to automation is not well examined.[1] The directional migration model focuses on gross population flows as opposed to net changes in the share or the net population growth of each occupation, thereby allowing researchers to examine the occupational composition change through migration more directly. The present study bridges this gap in the literature by analyzing how automation has influenced US interstate migration in the 21st century. In particular, using migration flow data classified into two skill groups and data classified into five major occupational groups, we examine how automation is changing the composition of local labor markets in the United States.

The previous literature on internal migration has studied the determinants and outcomes of migration on local labor markets (Greenwood, 1997; Molloy et al., 2011), finding that the

[1] Autor and Dorn (2013) also examine the effect of computerization on the difference in education shares between migrant and non-migrant workers, but do not focus on the direction of migration flows.

determinants include amenities, housing markets, and local labor market conditions (Davies et al., 2001; Greenwood and Hunt, 1989; Plantinga et al., 2013; Sasser, 2010; Zabel, 2012). Many empirical studies have focused on internal migration in the United States, where mobility has been high historically. However, mobility has started to decline since the 1980s, as explained by Molloy et al. (2011), Pingle (2007), and Modestino and Dennett (2013). In addition, several studies have focused on the migration of high-skilled workers whose human capital contributes to economic growth. For example, Fu and Gabriel (2012) find the importance of human capital concentration, especially of high-skilled migrants, for economic growth in destination regions in China. To the best of our knowledge, however, the effect of technological change in internal migration has not thus far been studied. This is the major contribution of the present study.

Following previous research on migration flows, we employ the conditional logit approach to examine migration decisions. This approach verifies whether people migrate to a state with a higher automation level than the origin state ceteris paribus. To measure the level of automation in each state, we calculate the employment share of highly automated occupations. The degree of automation for each occupation is obtained from the Occupational Information Network (O*NET), which is the primary occupational database in the United States. These valuable data, which have not thus far been used to study automation, provide direct information that is usually difficult to measure. This dataset thus enables us to consider how automation affects local labor markets more precisely.

In the first analysis, annual migration data at the interstate level from 2004 to 2016 are used to examine the extent to which automation affects migration flows. It is shown that people migrate to states with more automation in skilled occupations and less automation in unskilled occupations. In the second analysis, we use interstate migration flow data from 2004 to 2016 classified into several subgroups. We find that both skilled and unskilled workers move to states with more (less) automation in skilled (unskilled) occupations. The data classified into five occupational groups also show that most workers move to states with more (less) automation in skilled (unskilled) occupations. However, the effect of automation in unskilled occupations is less significant and it does not affect the migration decisions of those in particular occupations (production, transportation, and material moving occupations). Throughout the analyses, the results imply that automation in skilled (unskilled) occupations has a complementary (substitution) effect.[2]

The remainder of this chapter is organized as follows. Section 1.2 explains the literature on the economic theory of automation. Section 1.3 describes the model and Section 1.4 presents the data on the degree of automation by occupation. Section 1.5 reviews the data on migration flows and the other variables. Section 1.6 presents the results and Section 1.7 concludes.

1.2 Automation in local labor markets

This section discusses the economic theory underlying the potential effects of automation on migration decisions. Previous studies (particularly Autor et al. (2003)) argue that computer capital—as an instrument for automation—can replace routine work that follows explicit rules and is programmable. At the same time, computer capital is thought to complement non-

[2]Here, substitution and complementary effects are relationships between workers and computer capitals.

routine problem-solving and complex communication tasks. Autor et al. (2003) show that both these substitution and complementary effects have stimulated the demand for highly-educated workers with a comparative advantage in non-routine tasks. Autor and Dorn (2013) also show that middle-skilled occupations are more routine task-intensive and that computerization reallocates low-skill workers from routine task-intensive to service occupations, which are difficult to automate. This results in job polarization.

More specifically, Autor and Dorn (2013) and Accetturo et al. (2014) explain that job polarization occurs in the local labor markets where automation is taking place. Autor and Dorn (2013) show that in regions where production is intensive in routine tasks more computer capital is introduced as its price falls. Since computer capital also complements high-skilled workers, their wages in these regions rise faster than those of non-routine low-skilled workers. The authors conclude that high-skilled workers' welfare grows faster in these regions and that the region with the highest intensity of routine tasks in its production has the largest influx of high-skilled workers.[3] This finding suggests that regions characterized by a large number of automated jobs, especially in high-skilled occupations, have more in-migrants in those occupations. This results in agglomeration, which in turn increases the demand of these workers in service sectors. Consequently, these regions also witness an increase in workers in service jobs.

Automation may foster this population influx also because it produces new job opportunities. Berger and Frey (2016) show that, after the Computer Revolution in the 1980s, new jobs were created in US cities with endowments of analytical and interactive skills. The authors explain this phenomenon based on the complementarities between new technologies and skill endowments. A similar finding is presented by Gaspar and Glaeser (1998), although they focus on the effect of IT changes on face-to-face communication, showing that in some cases the latter has been replaced by electronic tools. However, when face-to-face communication is important at the point of contact, IT changes have actually a positive effect owing to the complementary effect. Automation may follow a similar pattern, overall suggesting that regions with more automation in skilled occupations attract many in-migrants. On the contrary, automation has a substitution effect especially on routine unskilled workers. It suggests that they move to states with less automation in unskilled or that they change their job from routine task-intensive occupations to service occupations in the same local labor markets.

1.3 Model

This study uses a conditional logit model (McFadden, 1973) following research on migration flows (Davies et al., 2001; Fu and Gabriel, 2012; Poncet, 2006; Sasser, 2010). We consider the situation in which an individual in origin region i has J choices: moving to destination region j or staying in origin region i. The individual chooses a region by maximizing his/her utility, which depends on the characteristics of the destination region and cost of moving from origin

[3] Autor and Dorn (2013) assume that only high-skilled workers can migrate across regions, whereas low-skilled workers cannot. However, they state that the similar result holds when low-skilled workers can also migrate.

region i to destination region j:

$$u_{ij} = \begin{cases} \beta' X_j + \epsilon_{ij} & if \quad j \neq i \ (moving) \\ \beta' X_i + \epsilon_{ii} & if \quad j = i \ (not\ moving) \end{cases}$$

where X_j is a vector of region j's features. β is the coefficients of X_j and ϵ_{ij} is an idiosyncratic error.

The probability of moving from region i to region j is

$$P_{ij} = \text{Prob}(u_{ij} > u_{ik}) \ \forall k \neq j$$

Based on McFadden (1973), if ϵ_{ij} is independent and identically distributed with a type-I extreme value distribution, the probability of moving from region i to region j can be given by

$$P_{ij} = \begin{cases} \frac{\exp(\beta' X_j)}{\sum_k \exp(\beta' X_k)} & if \quad j \neq i \ (moving) \\ \frac{\exp(\beta' X_i)}{\sum_k \exp(\beta' X_k)} & if \quad j = i \ (not\ moving) \end{cases}$$

Therefore, by taking the ratio of P_{ij} and P_{ii},

$$\frac{P_{ij}}{P_{ii}} = \frac{\exp(\beta' X_j)}{\exp(\beta' X_i)}$$
$$\Leftrightarrow \ \log(\frac{P_{ij}}{P_{ii}}) = \beta'(X_j - X_i)$$

In the analysis, X_j is the vector of the target independent variables, one of which is the employment share of highly automated occupations (termed the automated employment share herein, or AES), as explained below. In addition, X_j contains the control variables Z_j, which are the population, wage, housing price index, employment growth rate, and unemployment rate following (Davies et al., 2001; Sasser, 2010). Z_j also includes the share of high-skilled employment to total employment (Fu and Gabriel, 2012) and the employment share by industry. Each independent variable is the difference between the origin region and destination region for each of these variables. Further, the independent variables are lagged by one year to ease the endogeneity problem. The specification form used to analyze the effect of automation on migration decisions is

$$\log(\frac{P_{ijt}}{P_{iit}}) = \alpha + \beta(AES_{jt-1} - AES_{it-1}) + \delta(Z_{jt-1} - Z_{it-1}) + s_{ij} + q_t + e_{ijt} \tag{1.1}$$

where β and δ are the coefficients of the difference in the target independent variable and the differences in the control variables, respectively. α is the constant term. s_{ij} and q_t are the fixed effects of the state pair and year, respectively. The fixed effects of the state pair represents the time-invariant characteristics of each state pair such as the distance between them, which reflects migration costs, and idiosyncratic migrations such as pursuing college education or getting married. We assume that the bilateral relationship is symmetric, $s_{ij} = s_{ji}$. This chapter uses over-time variation within state pairs to examine whether and how the out-migration rate changes if the difference in the employment share of highly automated occupations between the states changes. e_{ijt} is the measurement and specification error. The standard errors are clustered by state pair. To distinguish the effect of automation at different skill levels, the following specification form includes $AES_{skilled}$ and $AES_{unskilled}$ instead of AES:

$$\begin{aligned}\log\left(\frac{P_{ijt}}{P_{iit}}\right) = \alpha &+ \beta_1(AES_{skilled,\ jt-1} - AES_{skilled,\ it-1}) \\ &+ \beta_2(AES_{unskilled,\ jt-1} - AES_{unskilled,\ it-1}) \\ &+ \delta(Z_{jt-1} - Z_{it-1}) + s_{ij} + q_t + e_{ijt} \qquad (1.2)\end{aligned}$$

1.4 Degree of automation

To specify which occupations have a high automation level, this study uses the degree of automation for each occupation from O*NET.[4] O*NET is the primary source of occupational information in the United States including data on the knowledge, skills, abilities, and tasks for each occupation. A new version of O*NET has been released periodically since 1998.[5]

The O*NET database reports the degree of automation for each occupation in the "Work Context" section, using an index from 1 to 5 (more automated occupations have a higher index value). During the study period, the mean of the degree is 2.17 and the standard deviation is 0.58. The max value of the degree is 4.14 (extruding, forming, pressing, and compacting machine setters, operators, and tenders) and the minimum is 1 (massage therapists).[6]

Previous studies find that the impact of automation on jobs differs depending on skill level. In particular, middle-skilled occupations who mainly engage in routine jobs are considered to be replaced by machines. This study therefore uses the typical education level of jobs as a proxy for the skill level involved.[7] Figure 1.1 shows the transition of the mean of the degree of automation for each education level.[8] The groups with the highest value are "High school diploma or equivalent" and "Associate's degree," which can be interpreted as middle-skill levels. This confirms that middle-skilled occupations are most affected by automation. "Bachelor's degree," which can be interpreted as a high-skill level, also has a large value among the groups. The low-skill level, "Less than high school," comes next to them and the high-skill levels, "Master's degree" and "Doctoral or professional degree," have lower values. Interestingly, the mean value for the highest skill group "Doctoral or professional degree" is increasing. Although the value varies by skill level, all are around 2 (i.e., "slightly automated" according to the five-point scale). However, whether automation brings about a substitution effect rather than a complementary

[4]The previous title of the database was the Dictionary of Occupational Titles, which has been used in the literature of automation such as Autor et al. (2003), Autor and Dorn (2013), and Berger and Frey (2016).

[5]Handel (2016) reports the detailed data collection method of O*NET.

[6]Here, the mean, standard deviation, max, and min of the degree of automation are calculated for the standard occupational classification (SOC)-level occupations in O*NET-SOC, especially the occupations used for the analyses in this study.

[7]The education level of each automated occupation can be identified using the "Typical education needed for entry" from the Occupational Projections and Training Data (Employment Projections: 2010-2020) published by the Bureau of Labor Statistics and complied/distributed by the National Crosswalk Service Center.

[8]Again, the values are calculated for the SOC-level occupations in O*NET-SOC, especially those used for the analyses in this study.

effect to each skill level job is unclear at this stage.

The original occupational code in O*NET (O*NET-SOC) is based on the classification used by federal statistical agencies. O*NET-SOC2010 has 1,110 occupations comprising SOC-level O*NET-SOC occupations and detailed O*NET-SOC occupations, which are a finer class of SOC-level occupations. This study focuses on SOC-level O*NET-SOC occupations. The information on each occupation's education level, Occupational Projections and Training Data (the Bureau of Labor Statistics, the National Crosswalk Service Center), is based on SOC2010. To define the skill level for all O*NET-SOC occupations in each year, they must thus correspond to O*NET-SOC2010 and thus to SOC2010.[9] This study uses only those O*NET-SOCs that have corresponding O*NET-SOCs in all years. Therefore, it excludes the newly established O*NET-SOCs following the three taxonomy changes during the study period. It also excludes the O*NET-SOCs that integrated or split following the taxonomy changes.[10]

In empirical studies such as Autor et al. (2003) and Autor and Dorn (2013), the degrees of routine work in each occupation are used to represent the possibility of each worker being affected by automation. O*NET also contains similar data on the degree of job routine, "Importance of repeating the same tasks," which is represented by an index from 1 to 5. The correlation of the degree of automation and "Importance of repeating the same tasks" in the sample period is 0.395, suggesting weak correlation between them. We infer that this is because the degree of automation in the dataset includes not only a substitution effect but also a complementary effect.

1.5 Data and variables

The state-to-state migration flows used in this study are derived from the mobility flow data of the American Community Survey (ACS). In our baseline analysis, we use the "State-to-State Migration Flows" from 2005 to 2017, which present the annual migration flows between the 50 states and District of Columbia.[11] The study period in the analysis is 2004 to 2016. By way of an additional occupation analysis, we use two migration flow data classified into skill groups or occupational groups in the same period. We construct these data using the ACS sample from 2005 to 2017 from the Integrated Public Use Microdata Series. This is a 1-in-100 national random sample of the population that allows us to estimate the interstate migration flow of each skill and occupational group.[12] There are five occupational groups based on the

[9] The version of the O*NET for each year is following: O*NET 5.0 (based on O*NET-SOC2000) for 2003, O*NET 6.0 (based on O*NET-SOC2000) for 2004, O*NET 8.0 (based on O*NET-SOC2000) for 2005, O*NET10.0 (based on O*NET-SOC2006) for 2006, O*NET12.0 (based on O*NET-SOC2006) for 2007, O*NET13.0 (based on O*NET-SOC2006) for 2008, O*NET14.0 (based on O*NET-SOC2009) for 2009, O*NET15.0 (based on O*NET-SOC2009) for 2010, O*NET16.0 (based on O*NET-SOC2010) for 2011, O*NET17.0 (based on O*NET-SOC2010) for 2012, O*NET18.0 (based on O*NET-SOC2010) for 2013, O*NET19.0 (based on O*NET-SOC2010) for 2014, and O*NET20.0 (based on O*NET-SOC2010) for 2015.

[10] We estimate the same specification model with AES, $AES_{skilled}$, and $AES_{unskilled}$ constructed using all occupations. Skill level is defined using the "required level of education" in O*NET for each year. This implementation does not change the sign or significance of the estimated coefficients in terms of automation, especially in the specification model in Section 1.6.1.

[11] For example, "State-to-State Migration Flows: 2007" reports migration flows from 2006 to 2007.

[12] We use "Migration status, 1 year (whether the person had changed residence since a reference point a year ago)," "State or country of residence 1 year ago", and "Person weight," which indicates how many people in the US population are represented by a given person to calculate the migration flows between each states and non-migrants for each state.

SOC system: (i) management, professional, and related occupations; (ii) service occupations; (iii) sales and office occupations; (iv) construction, extraction, and maintenance occupations; and (v) production, transportation, and material moving occupations. Moreover, there are two skill groups: (i) skilled workers (Bachelor's degree or more) and (ii) unskilled workers. We use individuals in the labor force who do not live in group quarters and who have a job from which they are temporarily absent (e.g., on vacation), working, or seeking a job within a specific reference week.

We infer that workers decide to migrate in response to the local labor market conditions. This study adopts states as the geographic unit because of the limitation of the data used to construct migration flows classified into skill or occupational groups, which is the ACS one-year sample. As mentioned in U.S. Census Bureau (2018), "for geographic areas with smaller populations, the ACS samples too few housing units to provide reliable single-year estimates". In addition, constructed migration flow data between metropolitan statistical areas (MSAs) or counties has a severe zero flow problem. According to Molloy et al. (2011), "state and county lines are often used to approximate local labor markets. Fortunately, both provide a reasonable proxy for inter-metropolitan migration" because interstate migrants have changed local labor markets such as MSAs. Since 60% to 70% of migrants across metropolitan areas changed states in five-year migration statistics from the Census and one-year migration statistics from the ACS (Molloy et al., 2011), interstate migration can be interpreted as a proxy for migration across local labor markets, although it does underestimate the number of people that move across local labor markets. Therefore, this study estimates the lower-bound effect of automation.

As noted earlier, to measure how each state has been affected by automation, we focus on the employment share of highly automated occupations (i.e., occupations with a high automation level). The automated employment share for each state i in year t, AES_{it}, is calculated as

$$AES_{it} = (\sum_{k=1}^{K} L_{ikt} \cdot 1[degree_{kt} > degree_{\hat{k}t}])(\sum_{k=1}^{K} L_{ikt})^{-1}$$

where L_{ikt} is employment in occupation k in state i in year t. This study uses the Occupational Employment Statistics from the Bureau of Labor Statistics for the employment data.[13] K includes all the target occupations regardless of their skill level.[14] $degree_{kt}$ is the degree of the automation of occupation k in year t. We rank $degree_{kt}$ in ascending order and select the 66th percentile as $degree_{\hat{k}t}$. $1[\cdot]$ is the indicator function, which takes 1 if the occupation's degree of automation is above the 66th percentile value in the year (i.e., the occupation is highly automated).

Previous research shows that automation has a different impact on workers depending on their skill levels, especially replacing middle-skilled workers and complementing high-skilled workers. To distinguish the effect of automation in these different skill-grouped occupations, the AES of skilled occupations ($AES_{skilled}$) and AES of unskilled occupations ($AES_{unskilled}$) are constructed. Occupations who need Bachelor's, Master's, and Doctoral or professional degrees as typical education level for entry into employment are regarded as skilled and those who

[13] The Occupational Employment Statistics is a semi-annual survey that estimates the number of jobs for SOC occupations in each state.

[14] The occupations with missing values in the degree of automation are excluded from K.

need other education levels are regarded as unskilled. Then, $AES_{skilled}$ and $AES_{unskilled}$ are calculated as follows:

$$AES_{skilled,it} = (\sum_{k_s=1}^{K_S} L_{ik_st} \cdot 1[degree_{k_st} > degree_{\hat{k_s},t})(\sum_{k_s=1}^{K_S} L_{ik_st})^{-1}$$

$$AES_{unskilled,it} = (\sum_{k_u=1}^{K_U} L_{ik_ut} \cdot 1[degree_{k_ut} > degree_{\hat{k_u},t}])(\sum_{k_u=1}^{K_U} L_{ik_ut})^{-1}$$

where K_S and K_U are the set of occupations identified as skilled and unskilled, respectively. Thus, K_S and K_U are the subsets of K. $degree_{\hat{k_s},t}$ and $degree_{\hat{k_u},t}$ are the 66th percentiles of $degree_{kt}$ among skilled/unskilled occupations. The degree of automation from O*NET shows the level of automation for each occupation when the data were released. Thus, AES, $AES_{skilled}$, and $AES_{unskilled}$ represent the employment shares of occupations with a high automation level as opposed to occupations with a large change in automation degree.

Following previous analyses of internal migration (Davies et al., 2001; Sasser, 2010; Fu and Gabriel, 2012), the control variables include economic factors that may affect migration decisions such as the population, unemployment rate, employment growth rate, housing price index, and wages. State population measures the size of the economy. The population may also reflect the quantity and quality of all opportunities across states (Davies et al., 2001) because a more populated region can enjoy a higher variety of goods and services. Thus, people may have an incentive to move to more populated states. The unemployment rate and employment growth rate reflect the labor market conditions in the state. We conjecture that people have an incentive to migrate to a state with better labor market conditions (i.e., a lower unemployment rate and higher employment growth rate). The cost of living is represented by the housing price index. A higher cost of living and lower cost of wages[15] may increase the incentive to migrate. It could also be possible that migration flows are affected by the difference in the state's employment mix. In order to account for this, we use the SIC division to classify industries and include the difference in the employment shares of each industry as controls in the model.[16] Tables 1.1 and 1.2 report the sample statistics of the dependent and independent variables used in the analyses.

1.6 Impact of automation on migration: Regression results

1.6.1 State-to-state migration flows: Baseline results

Table 1.3 reports the regression results, showing the extent to which automation affects migration flows from 2004 to 2016 for the eight specifications. Columns 1 and 2 show only the main independent variables; the difference in AES is in column 1, while $AES_{skilled}$ and $AES_{unskilled}$

[15] This study uses the weekly average wage.

[16] The data sources of the control variables are as follows. Population: 2000–2010 Intercensal Estimates and State Population Totals Tables 2010–2016 (the Bureau of the Census). Wage: Quarterly Census of Employment and Wages (the Bureau of Labor Statistics). Unemployment rate and employment growth rate: Local Area Unemployment Statistics (the Bureau of Labor Statistics). House price index: House Price Index Datasets (Federal Housing Finance Agency). Share of high-skilled workers: Occupational Employment Statistics (the Bureau of Labor Statistics) and Occupational Projections and Training Data (the Bureau of Labor Statistics, the National Crosswalk Service Center). Employment share by industry: Gross Domestic Product by State (the Bureau of Economic Analysis).

are in column 2. Columns 3 and 4 add the control variables to columns 1 and 2 with no fixed effects. Columns 5 and 6 include all the independent variables and fixed effects for the year and state pair. Finally, columns 7 and 8 add an additional control variable, the state's difference in the rate of unionization, to columns 5 and 6. To enhance our understanding of the impact of economic factors on migration flows, the independent variables are standardized to have a zero mean and one standard deviation. One point to be noted is the zero-flow problem, which is similar to the zero-trade problem in gravity estimation. Several destination–origin pairs have no migrants. Of the 33,150 observations, 2,445 have zero flows. Since the dependent variable in the analysis is the logarithm of the ratio of the flow to non-migrants, the dependent variables for the zero-flow observations are undefined. We thus omit the observations with zero flows and run a classical OLS.[17]

To examine the extent to which automation affects migration decisions, columns 1, 3, and 5 use the difference in AES as the main independent variable. The results show that the difference in AES is positively significant without the control variables in column 1. However, once we include the control variables in columns 3 and 5, the estimated coefficient becomes insignificant. Based on the results with the controls, automation does not affect migration. Here, highly automated occupations include both skilled and unskilled occupations. Since about 80% of highly automated occupations are unskilled occupations in the study period, this result therefore captures the effect of automation in most unskilled occupations.

In columns 2, 4, and 6, we distinguish between the effect of automation in skilled and unskilled occupations. Thus, the differences in $AES_{skilled}$ and $AES_{unskilled}$ are included as the main independent variables instead of AES. According to economic theory, automation replaces unskilled workers and complements skilled workers. We can thus conjecture that people migrate to a state with low $AES_{unskilled}$ and high $AES_{skilled}$. When we do not include the controls, the coefficients of the differences in both $AES_{skilled}$ and $AES_{unskilled}$ are positively significant. Once we include the controls, however, while the difference in $AES_{skilled}$ is still positively significant, the difference in $AES_{unskilled}$ becomes negative and insignificant. This result implies that automation in skilled occupations has a complementary effect, or at least not a strong substitution effect. On the contrary, automation in unskilled occupations has an insignificant substitution effect, or at least not a strong complementary effect. The absolute value of the coefficient of the difference in $AES_{skilled}$ is larger than that of the difference in $AES_{skilled}$. This finding implies that the complementary effects of automation in skilled occupations are stronger than the substitution effects of automation in unskilled occupations.

Columns 7 and 8 include the difference in the rate of unionization as an additional control. Unionization often hinders automation, especially in unskilled occupations, since such jobs are threatened to be replaced by automation. Because unionization may affect the automated employment share and migration flows, we use the state's percentage of union members based on Labor Force Statistics from the Current Population Survey (the Bureau of Labor Statistics). When we include the controls, the absolute value of the estimated coefficient of $AES_{skilled}$ and $AES_{unskilled}$ become smaller and larger, respectively. Moreover, $AES_{unskilled}$ becomes significant. This finding implies that when controlling for the effect of unionization, the complementary

[17] We also estimate the same specification model with zero flows. The dependent variable is redefined by $\log((P_{ij}+1)/(P_{ii}+1))$. This implementation does not change the conclusions of the analyses.

effect of automation in skilled occupations weakens and the substitution effect of automation in unskilled occupations strengthens.

For the other economic factors, the size and sign of the coefficients coincide in most of the specifications and the results are consistent with our expectations. Out-migration is higher when the housing price index is high in the origin state relative to in the destination state. That is, people migrate to a state with a lower cost of living. In terms of the population, people migrate to more populated states. The controls representing local economic conditions (the difference in employment growth rate, and unemployment rate) are also consistent with the expectations that people move to a state with better conditions. The coefficients of the difference in the population are the largest among the independent variables. The housing price index is the second largest, which agrees with the recent trend that the importance of housing affordability on migration decisions has risen (Sasser, 2010). Wages have a small or insignificant impact on migration, which again concurs with the tendency of per-capita income to have fallen since the late 1970s (Sasser, 2010).

1.6.2 State-to-state migration flows by occupation and skill

To analyze who is particularly influenced by automation, we use two migration flow datasets classified into subgroups. The first dataset is classified into two skill groups: (i) skilled workers and (ii) unskilled workers. The other dataset is classified into five occupational groups: (i) management, professional, and related occupations; (ii) service occupations; (iii) sales and office occupations; (iv) construction, extraction and maintenance occupations; and (v) production, transportation, and material moving occupations. Of these, about 70% of the occupations in category (i) are skilled and about 95% of the occupations in each categories (ii)–(v) are unskilled. The independent variables are again scaled to have zero means and one standard deviation. As with the analysis of overall interstate migration, we omit observations with zero flows.

The results in Table 1.4 show that both skilled and unskilled workers have positive significance in the difference in $AES_{skilled}$, implying that people move to a state with more automation in skilled occupations irrespective of their skill. On the contrary, the difference in $AES_{unskilled}$ is negative for both workers, implying that both skill workers move to states with less automation in unskilled occupations. Table 1.5 shows the simulation results of calculating how many workers in each skill group move with a one standard deviation increase in the difference in $AES_{skilled}$ and $AES_{unskilled}$ following Sasser (2010). It shows that unskilled workers move more than skilled workers in response to an increase in the difference in $AES_{skilled}$, whereas skilled workers move more than unskilled workers in response to an increase in the difference in $AES_{unskilled}$. Thus, automation in both skill occupations not only has complementary/substitution effects on their workers but also has spillover effects onto workers in a different skill occupation.

According to the results based on the migration flow data classified into occupational groups in Table 1.4, all occupations have positive coefficients of the difference in $AES_{skilled}$ and negative coefficients of the difference in $AES_{unskilled}$. However, their significance levels vary. The difference in $AES_{skilled}$ is highly significant for all occupations but occupation (v) (production, transportation, and material moving occupations) which has less significance. In terms of the difference in $AES_{unskilled}$, its high significance is limited to particular occupations: occupation (i) (management, professional, and related occupations), occupation (ii) (service occupations) and occupation (iv) (construction, extraction and maintenance occupations). The results also show that automation in unskilled occupations does not affect the migration decision of workers in occupation (v) (production, transportation, and material moving occupations). The simulation results in Table 1.6 show that the major migrants for automation both in skilled and in unskilled jobs are in occupation (i) (management, professional, and related occupations) and the next major migrants are in occupation (iii) (sales and office occupations). Occupation (ii) (service occupations), which seems irrelevant to any production technology change, also responds to automation in skilled work. Here, the productivity growth brought about by automation in skilled work increases demand for service goods, as suggested by studies of job polarization such as Autor and Dorn (2013).

1.7 Conclusion

The development of AI has focused attention on the fear of people being replaced by computers or robots in the workplace. However, automation is thought to have a complementary effect on skilled occupations that raises productivity. Given the substitution and complementary effects of automation, we examined which is stronger in local labor markets as well as whether automation will change the population and occupational distributions among states. In particular, using data on the degree of automation for each occupation from O*NET, the level of automation in each state was calculated. We then analyzed how the difference in the levels of automation among states affects interstate migration flows using the conditional logit approach as well as how automation will change the occupational distribution among states.

The results showed that people migrate to both states with more automation in skilled occupations and states with less automation in unskilled occupations. This finding implies that automation in skilled occupations has a complementary effect, whereas automation in unskilled occupations has a substitution effect. Further, the analyses with migration flow data classified into skill or occupational groups showed that the complementary effect of automation in skill occupations is highly significant and affects the migration decisions of both skilled and unskilled workers in all occupations. However, the substitution effect of automation in unskilled occupations is less significant and it does not affect the migration decisions of those in particular occupations (production, transportation, and material moving occupations).

Since a state with more automation in skilled occupations attracts both skilled and unskilled migrants, this type of automation will deliver the state economic growth. Automation in both skill occupations not only has direct effects on their own workers but also has spillover effects onto workers in a different skill occupation. For example, the second major category of migrants in response to automation in skilled work is workers in sales and office occupations, which may be classified into unskilled occupations, but are related to skilled occupations. In addition, mobility is high in service occupations, which seems to be irrelevant to technological change. This is consistent with the findings of Autor and Dorn (2013), who show that technological change increases demand for jobs with manual tasks. Surprisingly, automation in unskilled occupations has spillover effects onto workers in skilled occupations and they are the major migrants for automation in unskilled. This finding implies that skilled workers move from states with more automation in unskilled work, while unskilled workers stack in there, which hinders a state's economic growth.

Lastly, one limitation of this study should be noted. We assumed that workers do not change jobs when they migrate, especially in the analyses with migration flow data classified into occupational groups. However, when they move to seek a new job, they may be influenced by automation, especially if their jobs are being replaced with robots or computers. Changing this assumption to consider the job change effect is thus suggested as a task for future research.

Figure 1.1: Mean of the degree of automation for each education level

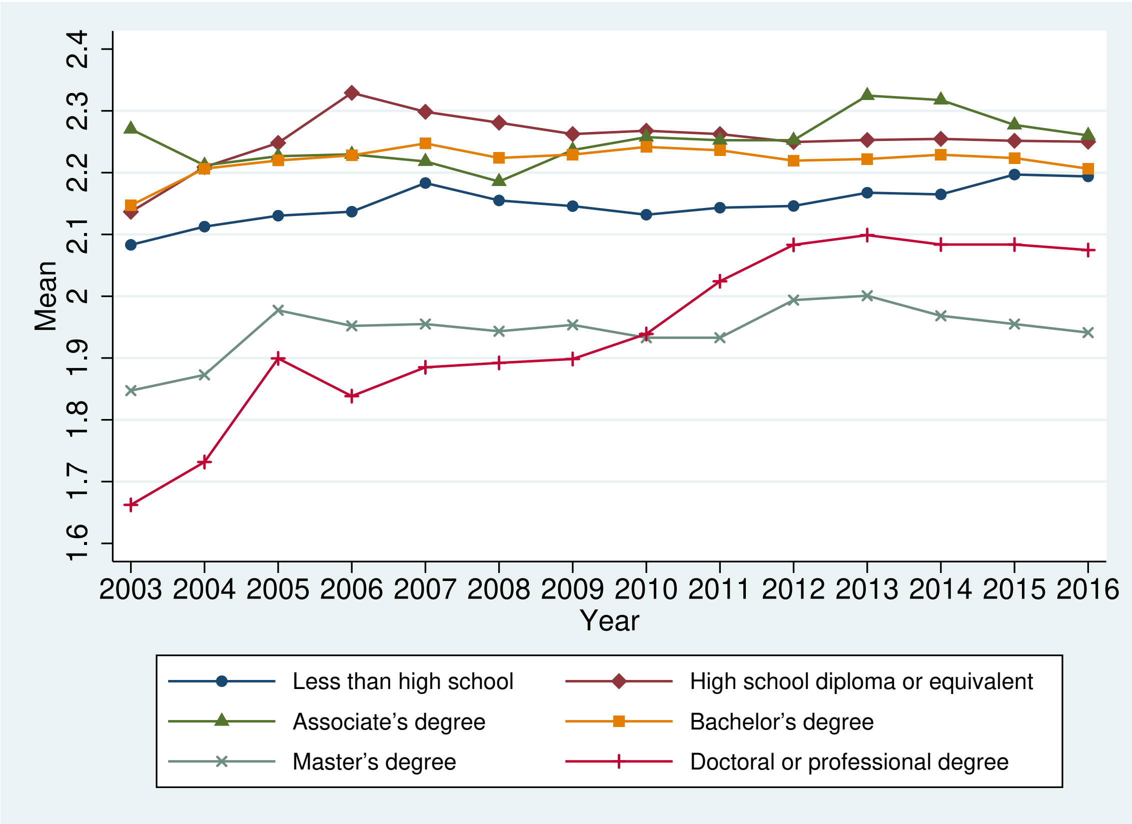


Table 1.1: Summary statistics of the dependent variables

	Number of out-migrants		Number of non-migrants		Ratio of out-migrants to non-migrants	
	Mean	S.D	Mean	S.D	Mean	S.D
All	3078.9	5883	781314	894283.1	.0056	.0122
Occup.1: Management, professional, and related occupations	880.5	1401.1	1248522	1282412	.0011	.0019
Occup.2: Service occupations	489.5	685.7	709387.9	726637.3	.0013	.0027
Occup.3: Sales and office occupations	601.8	909.3	943806.4	953798.4	.0011	.0024
Occup.4: Construction, extraction and maintenance occupations	330.2	417.1	382548.1	374320.1	.0016	.0035
Occup.5: Production, transportation and material moving occupations	365	455.8	517407.2	471917.5	.0014	.0032
Skilled workers	927.3	1495.3	1102966	1163044	.0013	.0022
Unskilled workers	1189.1	2049.7	2461474	2531078	.0008	.0019

Table 1.2: Summary statistics of the independent variables

	Mean	S.D.	Min	Max
AES	0.32	0.03	0.25	0.43
$AES_{skilled}$	0.39	0.06	0.26	0.63
$AES_{unskilled}$	0.26	0.05	0.10	0.39
Housing price index	138.02	25.60	83.02	267.65
Wage (USD)	825.02	172.54	517.00	1695.00
Share of high-skilled workers	0.15	0.03	0.08	0.35
Employment growth rate	0.01	0.02	-0.07	0.05
Unemployment rate	0.06	0.02	0.03	0.14
Population (in thousands)	6002.39	6718.37	503.45	38993.94
Rate of unionization	0.11	0.05	0.02	0.26
Employment share by industry: Agriculture, Forestry, and Fishing	0.02	0.02	0.00	0.13
Employment share by industry: Construction	0.04	0.01	0.01	0.11
Employment share by industry: Finance, Insurance, and Real Estate	0.19	0.05	0.08	0.47
Employment share by industry: Manufacturing	0.12	0.06	0.00	0.31
Employment share by industry: Mining	0.03	0.07	0.00	0.39
Employment share by industry: Public Administration	0.14	0.04	0.09	0.37
Employment share by industry: Retail Trade	0.09	0.03	0.04	0.23
Employment share by industry: Servicies	0.24	0.05	0.10	0.46
Employment share by industry: Transportation, Communications and Public Utilities	0.05	0.02	0.01	0.13
Employment share by industry: Wholesale Trade	0.06	0.01	0.01	0.08
		Absolute value		
	Mean	S.D.	Min	Max
Diff. AES	0.76	0.64	0.00	4.60
Diff. $AES_{skilled}$	0.72	0.69	0.00	5.27
Diff. $AES_{unskilled}$	0.75	0.66	0.00	6.31
Diff. housing price index	0.74	0.68	0.00	5.08
Diff. employment growth rate	0.79	0.62	0.00	3.91
Diff. unemployment rate	0.77	0.63	0.00	4.80
Diff. population	0.65	0.76	0.00	4.00
Diff. share of high-skilled workers	0.69	0.72	0.00	4.91
Diff. wage	0.72	0.69	0.00	4.48
Diff. rate of unionization	0.80	0.60	0.00	3.08
Diff. employment share by industry: Agriculture, Forestry, and Fishing	0.02	0.02	0.00	0.13
Diff. employment share by industry: Construction	0.01	0.01	0.00	0.09
Diff. employment share by industry: Finance, Insurance, and Real Estate	0.06	0.05	0.00	0.35
Diff. employment share by industry: Manufacturing	0.07	0.05	0.00	0.30
Diff. employment share by industry: Mining	0.05	0.08	0.00	0.39
Diff. employment share by industry: Public Administration	0.04	0.05	0.00	0.27
Diff. employment share by industry: Retail Trade	0.02	0.03	0.00	0.19
Diff. employment share by industry: Services	0.06	0.05	0.00	0.37
Diff. employment share by industry: Transportation, Communications and Public Utilities	0.02	0.02	0.00	0.11
Diff. employment share by industry: Wholesale Trade	0.01	0.01	0.00	0.08

The difference variables are calculated as the absolute standardized value of difference between the value in the destination state minus the value in the origin state. The wage is the annual average weekly wage.

Table 1.3: Estimation of the relationship between the economic conditions and migration flows for 2004–2016

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Diff. AES	0.150***		0.005		0.007		-0.013	
	(0.020)		(0.010)		(0.010)		(0.009)	
Diff. AES skilled		0.152***		0.076***		0.080***		0.074***
		(0.021)		(0.008)		(0.008)		(0.007)
Diff. AES unskilled		0.119***		-0.007		-0.007		-0.027***
		(0.019)		(0.011)		(0.010)		(0.010)
Diff. housing price index			-0.259***	-0.246***	-0.258***	-0.245***	-0.239***	-0.227***
			(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
Diff. employment growth rate			0.045***	0.039***	0.048***	0.041***	0.035***	0.029***
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Diff. unemployment rate			-0.011	-0.007	-0.013	-0.008	-0.001	0.002
			(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Diff. population			0.529***	0.528***	0.531***	0.529***	0.531***	0.530***
			(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)
Diff. share of high-skilled workers			0.087***	0.096***	0.089***	0.100***	0.077**	0.089***
			(0.032)	(0.032)	(0.031)	(0.031)	(0.031)	(0.031)
Diff. wage			-0.004	-0.048*	-0.011	-0.057**	0.028	-0.021
			(0.026)	(0.026)	(0.025)	(0.025)	(0.026)	(0.026)
Diff. rate of unionization							-0.092***	-0.089***
							(0.012)	(0.011)
Diff. employment share by industry			✓	✓	✓	✓	✓	✓
State-pair and year fixed effects					✓	✓	✓	✓
N	30705	30705	30705	30705	30705	30705	30705	30705
R^2	0.010	0.018	0.264	0.265	0.754	0.756	0.756	0.758
adj. R^2	0.010	0.018	0.263	0.265	0.743	0.745	0.745	0.747

*p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered by state pair. The independent variables are lagged by one year and are calculated as the difference between the value in the destination state minus the value in the origin state. The independent variables are standardized to have zero means and one standard deviation. The SIC division is used as the industry classification. Thus, there are 10 controls in terms of the difference in employment share by industry.

Table 1.4: Estimation of the relationship between the economic conditions and migration flows for 2004–2016

	Skill group:		*Occupational group:*				
	Skilled	Unskilled	Occup. 1	Occup. 2	Occup. 3	Occup. 4	Occup. 5
Diff. AES skilled	0.047***	0.055***	0.045***	0.033***	0.042***	0.046***	0.030**
	(0.009)	(0.009)	(0.009)	(0.012)	(0.010)	(0.013)	(0.013)
Diff. AES unskilled	-0.062***	-0.029***	-0.054***	-0.043***	-0.033**	-0.049***	-0.001
	(0.011)	(0.011)	(0.011)	(0.014)	(0.013)	(0.018)	(0.018)
Diff. housing price index	-0.122***	-0.165***	-0.134***	-0.163***	-0.131***	-0.123***	-0.173***
	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.021)	(0.020)
Diff. employment growth rate	0.005	0.007	0.005	-0.000	0.009	0.004	0.000
	(0.007)	(0.007)	(0.007)	(0.009)	(0.008)	(0.011)	(0.011)
Diff. unemployment rate	-0.058***	-0.052***	-0.064***	-0.066***	-0.033**	-0.064***	-0.078***
	(0.012)	(0.013)	(0.013)	(0.015)	(0.013)	(0.019)	(0.017)
Diff. population	0.468***	0.464***	0.471***	0.444***	0.461***	0.395***	0.393***
	(0.011)	(0.012)	(0.011)	(0.013)	(0.012)	(0.015)	(0.014)
Diff. share of high-skilled workers	0.209***	0.085***	0.189***	0.059	0.114***	0.027	-0.010
	(0.032)	(0.032)	(0.033)	(0.039)	(0.038)	(0.047)	(0.040)
Diff. wage	-0.022	-0.086***	-0.055*	-0.042	-0.088**	-0.043	-0.087*
	(0.028)	(0.030)	(0.029)	(0.038)	(0.036)	(0.055)	(0.050)
Diff. rate of unionization	-0.025*	-0.053***	-0.036***	-0.020	-0.048***	-0.076***	0.006
	(0.013)	(0.012)	(0.013)	(0.015)	(0.013)	(0.019)	(0.017)
Diff. employment share by industry	✓	✓	✓	✓	✓	✓	✓
State-pair and year fixed effects	✓	✓	✓	✓	✓	✓	✓
N	22187	23118	21903	16091	17940	11105	12494
R^2	0.674	0.676	0.668	0.618	0.660	0.567	0.589
adj. R^2	0.655	0.657	0.647	0.588	0.635	0.522	0.550

*p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered by state pair. The independent variables are lagged by one year and are calculated as the difference between the value in the destination state minus the value in the origin state. The independent variables are standardized to have zero means and one standard deviation. The SIC division is used as the industry classification. Thus, there are 10 controls in terms of the difference in employment share by industry. The dataset is classified into two skill groups: skilled workers (those with a Bachelor's degree or more) and unskilled workers (those without a Bachelor's degree). It is also classified into five occupational groups: (1) management, professional, and related occupations; (2) service occupations; (3) sales and office occupations; (4) construction, extraction and maintenance occupations; and (5) production, transportation, and material moving occupations.

Table 1.5: Simulation of the impact of a one standard deviation change in the difference in $AES_{skilled}$ and $AES_{unskilled}$ on skill group

	Skill group	(a) Mean of out-migration rate	(b) Mean of number of non-migrants	(c) Percent change in out-migration rate for a 1 SD increase in differential	(d) New implied out-migration rate for a 1 SD increase in differential	(e) Implied change in number of out-migrants for a 1 SD increase in differential
$AES_{skilled}$	Skilled workers	0.00130	1102966.4	0.04740***	0.00137	68.2
	Unskilled workers	0.00077	2461474.0	0.05470***	0.00081	103.9
$AES_{unskilled}$	Skilled workers	0.00130	1102966.4	-0.06192***	0.00122	-89.1
	Unskilled workers	0.00077	2461474.0	-0.02933***	0.00075	-55.7

(a) and (b) are taken from Table 1.1. (c) are the regression coefficients of $AES_{skilled}$ or $AES_{unskilled}$ in Table 1.4. (d) are calculated by (a) times 1 + (c). (e) are calculated by ((d) - (a)) times (b).

Table 1.6: Simulation of the impact of a one standard deviation change in the difference in $AES_{skilled}$ and $AES_{unskilled}$ on occupational group

	Occup. group	(a) Mean of out-migration rate	(b) Mean of number of non-migrants	(c) Percent change in out-migration rate for a 1 SD increase in differential	(d) New implied out-migration rate for a 1 SD increase in differential	(e) Implied change in number of out-migrants for a 1 SD increase in differential
$AES_{skilled}$	Occup. 1	0.00107	1248522	0.04487***	0.00112	60.1
	Occup. 2	0.00128	709387.9	0.03330***	0.00132	30.3
	Occup. 3	0.00113	943806.4	0.04197***	0.00118	44.7
	Occup. 4	0.00163	382548.1	0.04638***	0.00171	28.9
	Occup. 5	0.00141	517407.2	0.02971**	0.00145	21.7
$AES_{unskilled}$	Occup. 1	0.00107	1248522	-0.05356***	0.00102	-71.8
	Occup. 2	0.00128	709387.9	-0.04264***	0.00123	-38.8
	Occup. 3	0.00113	943806.4	-0.03329**	0.00109	-35.4
	Occup. 4	0.00163	382548.1	-0.04898***	0.00155	-30.6
	Occup. 5	0.00141	517407.2	-0.00146	0.00141	-1.1

(a) and (b) are taken from Table 1.1. (c) are the regression coefficients of $AES_{skilled}$ or $AES_{unskilled}$ in Table 1.4. (d) are calculated by (a) times 1 + (c). (e) are calculated by ((d) - (a)) times (b).

Chapter 2

Impacts of high-speed rail construction on urban agglomerations: Evidence from Kyushu in Japan

2.1 Introduction

Economic integration is no doubt a major factor shaping economic geography. In addition, railroads have also undoubtedly played a significant role in shaping the economic landscape. In fact, spatial economics literature has often focused on the impacts of decreases in transportation costs on economic geography, and such costs have been lowered by construction of railroad networks. Krugman (1991, page 487, l.2-l.6) referred to this point as "*...let the factory system and eventually mass production emerge, and with them economies of large-scale production; and let canals, railroads, and finally automobiles lower transportation costs. Then, the tie of production to the distribution of land will be broken...*" In fact, railroad construction can sometimes change economic geography even over megalopolises. We can find such an example in Japan: it is widely believed that the construction of a high-speed rail, Shinkansen, between the two largest cities in Japan, Tokyo and Osaka, attracted economic activities to Tokyo and caused Osaka's economy to decline.[1] Until 1960s, Osaka had been Tokyo's rival and had served as the hub of western Japan. However, after the Shinkansen's opening in 1964, many firms shifted their headquarters and substantial amounts of significant functions from Osaka to Tokyo, as indicated by steady outflows of population from the Osaka metropolitan area during the last half century. At present, the new Linear Shinkansen connecting Tokyo and Nagoya, which is the third largest city, is under construction and it will open by 2027. This is expected to reduce the travel time between the two cities from 150 minutes to 70 minutes. How and to what degree do such high-speed railways affect the economic geography?

This chapter aims to provide clues to answer these questions by quantifying the effects on urban agglomerations of the opening and extension of the high-speed rail in Kyushu, Japan's third largest island. The high-speed rail in Kyushu, called the Kyushu Shinkansen, currently operates in western Kyushu and is called the Kagoshima route. Its southern half opened in 2004 and was extended to the northern half in 2011 to traverse Kyushu longitudinally. Thus, the Kagoshima route of the Kyushu Shinkansen has significantly integrated Kyushu's regions. We examine its impacts on urban agglomerations by quantifying the changes capitalized in land prices. Economic integration would influence various economic activities from communication among people to goods and service trade. It would of course be significant to examine the impact of integration on each activity one by one. However, in this chapter, we focus on the people's overall evaluation. One way to capture it to examine the land price, in which values arising from various economic activities are capitalized. Hence, we estimate hedonic price

[1] See, for example, the PRESIDENT online article (October 30, 2017) by Kenichi Omae: http://president.jp/articles/-/23444 (accessed on February 2, 2018).

equations and conduct a difference-in-differences (DID) analysis. We use Urban Employment Areas (UEAs) as the definition of urban agglomerations; UEAs are defined similarly to the Consolidated Metropolitan Statistical Areas and are made from Japanese municipalities (see Kanemoto and Tokuoka, 2002). We construct panel data by using land prices and attributes of locations in Kyushu for 3 years before and after the railway's opening and extension. We then estimate how the opening and extension of the Kagoshima route changed land prices in urban agglomerations.

The estimation results show that the railway's opening and extension have increased land prices greatly in the large metropolitan areas and slightly in their neighboring areas: the opening of the entire Kagoshima route increased land prices in the largest metropolitan area in Kyushu, Fukuoka, by around 10.5%, and prices in the second largest metropolitan area, Kumamoto, by around 7.7%. However, if we extend the estimation by categorizing locations into different groups according to distance to the nearest Shinkansen station, we find such positive effects are limited to areas close to Shinkansen stations where land prices were already higher before the opening. In contrast to these rises, we further find that small metropolitan areas located between large metropolitan areas experienced decreases in land prices. Thus, integration caused by the Kagoshima route has made already-large metropolitan areas larger at the expense of small metropolitan areas and accelerated concentration also within enlarged metropolitan areas.

We can interpret our results by consulting the literature of New Economic Geography, which has investigated the roles of agglomeration and dispersion forces in shaping the economic geography. Its pioneering work, Krugman (1991), showed that integration between regions primarily accelerates agglomeration of economic activities to a particular region, resulting in the Core-Periphery structure. However, as shown by Helpman (1998) and Tabuchi (1998), if the dispersion forces include congestion diseconomies, a sufficient level of integration makes economic activities dispersed over regions. More recently, Akamatsu et al. (2017) showed that changes in economic geography caused by economic integration crucially depend on the characteristics of the dispersion force. When dispersion occurs globally through attraction from outside the agglomeration, as in the case of transporting goods and service to a distant, less-crowded market, integration enlarges already-larger agglomerations at the expense of smaller agglomerations. Dispersion may also occur locally to avoid crowding inside the agglomeration, as in the case of urban congestion diseconomies. In such circumstances, integration attenuates regional agglomeration and makes economic activities dispersed across the integrated regions. Moreover, economic integration affects distribution of economic activities within each region. When the global dispersion force is prominent, integration accelerates concentration within a region whereas when the local dispersion force is dominant, integration causes dispersion within a region.

Hence, it is possible that a region attracting economic activities due to integration has particular suburban areas that lose economic activities whereas it is also possible that a region losing economic activities due to integration has particular suburban areas that gains economic activities. This implies the needs to examine the effects location by location within a city/region. Our results show that integration increased land prices in already larger urban agglomerations and in central areas within each of them, implying that they are in line with the dominance of the global dispersion force.[2]

[2]Note here that our analysis is not a direct test of a particular model. Here we only claim that our results are

Our analysis is related to the literature studying the effect of construction of transportation infrastructure on the location of economic activities. The recent survey of the literature by Redding and Turner (2015) categorized existing studies into two groups: those regarding intracity transportation and those regarding intercity transportation. Our study belongs to the latter category wherein existing studies focused on various factors and reached different conclusions. For instance, a large strand of the literature found positive impacts of transportation infrastructure on cities/regions connected by it, although some studies found negative impacts of inter-city transportation infrastructure.[3] [4]

Such difference in conclusions would stem from possible regional heterogeneity in impacts of transportation infrastructure. We know from the literature of New Economic Geography that economic integration can have different impacts on heterogeneous cities/regions, which implies the needs to consider heterogeneity among cities/regions. For example, Faber (2014) found a county with larger market size and higher trade costs were less affected by highway construction, and Baum-Snow et al. (2016) found the primary city of a wider-region gained from the expansions of regional highway networks, of which the average effect on the wider-region was negative. Qin (2017) also showed that counties with highway access prior to the introduction of high-speed rail have less negative effects from the high-speed rail. Baum-Snow et al. (2018), which would be the most related paper to ours, showed that central cities gained economic outputs and population at the sacrifice of hinterland cities from the investments in national highways in China. It is the first paper to provide "econometric evidence for an 'urban hierarchy' at the regional level" (Baum-Snow et al. (2018)).

Moreover, as we discussed above, we should also take locational difference within a city/region into consideration. In this respect, we here raise Mori and Takeda (2018) as existing works that uncovered the effects of inter-city transport network on different locations within a city. We also address within-city locational difference in the effects of high-speed rail construction.

In comparison to these existing studies, our study has three distinct features. First, we focus on the land price, which is an aggregate of various factors that determine economic agents' location decisions, implying that our analysis can capture the overall effects of construction of high-speed rail on the distribution of economic activities. Second, we consider locational heterogeneity within each city as well as heterogeneity among cities in the treatment group. This enables us to relate our results to the theoretical literature on the New Economic Geography. Finally, the area of our focus, Kyushu has desirable geographic features for program evaluation because the treatment group (here, locations mostly in the western part of Kyushu) and control group (here, locations mostly in the eastern part of Kyushu) had a common trend in land prices

consistent with the predictions made by a New Economic Geography model having the dominant global dispersion force.

[3]See Ahlfeldt and Feddersen (2018), Atack et al. (2010), Audretsch et al. (2017), Banerjee et al. (2012), Berger and Enflo (2017), Chandra and Thompson (2000), Duranton and Turner (2012), **?**, Hornung (2015), Jedweb and Moradi (2016), Lin (2017), Michaels (2008), Mori and Takeda (2018) and Storeygard (2016) on regional outcomes (e.g., population, employment, employment share and GDP); or Donaldson (2018), Donaldson and Hornbeck (2016), Duranton et al. (2013) and Volpe Martincus and Blyde (2013) on trade. Sanchis-Guarner (2013) and Heurermann and Schmieder (2014) use worker-level datas and Bernard et al. (2019), Datta (2012), Ghani et al. (2016) and Gibbons et al. (2019) use firm-level datas to see more precise effects on them. Among above, Bernard et al. (2019) ,Heurermann and Schmieder (2014) and Lin (2017) focus on high-speed rail.

[4]See Faber (2014), Baum-Snow et al. (2016) and Qin (2017) on regional output. Qin (2017) focuses on high-speed rail.

before the event. In addition, SUTVA (Stable Unit Treatment Value Assumption) is reasonably satisfied because the two groups are divided by mountains.[5] These features of Kyushu enable us to identify total treatment effects in the western part of Kyushu and analyze heterogeneous treatment effects among agglomerations in the treatment group.

The remainder of this chapter is structured as follows. Section 2.2 describes the details of Kyushu Shinkansen and research background. Section 2.3 explains our data and section 2.4 describes our estimation strategies. Section 2.5 provides the estimation results and section 2.6 discusses econometric issues. Section 2.7 concludes.

2.2 Research Background

We first explain the Kagoshima route of Kyushu Shinkansen, which is the focus of this chapter, and discuss its possible impacts on urban agglomerations. The Kyushu Shinkansen is a high-speed railway, and its Kagoshima route connects Hakata and Kagoshima-Chuo Stations, which are located in the northern and southern part of Kyushu, respectively. It runs through the western part of Kyushu, with a length of 256.8 km and stopping at 12 twelve stations (Figure 2.1).

[Figure 2.1]

The Kyushu-Shinkansen was constructed and is owned by an independent administrative agency named the Japan Railway Construction, Transport and Technology Agency, and is operated by the Kyushu Railway Company.[6] As explained by Ministry of Land, Infrastructure and Tourism, the Kagoshima route project was launched as one of five Shinkansen projects in 1973 based on the Nationwide Shinkansen Railway Development Act. The prerequisites for the project to start were securing a budget for construction, profitability, investment efficiency, and agreements of the Kyushu Railway Company and jurisdictions on the Shinkansen line, all of which were carefully investigated before breaking ground. Construction of the line started in 1991. The southern part of the line (between Shin-Yatsushiro and Kagoshima-Chuo Stations) started operating in March 2004 and the remaining part (between Hakata and Shin-Yatsushiro Stations) commenced operations in March 2011. The former reduced the travel time between Hakata and Kagoshima-Chuo Stations from 3 hours and 40 minutes to 2 hours and 12 minutes, and the latter shortened it to 1 hour and 19 minutes. Overall, the construction of the Kagoshima route decreased the travel time between the two stations by nearly two-thirds.

In this chapter, we estimate the effects of the construction of the Kagoshima route of the Kyushu Shinkansen on the local economy, especially on land prices in urban agglomerations in Kyushu, by using a hedonic approach a la Rosen (1974). In so doing, we conduct a difference-in-difference estimation using the data for the years 2001 and 2007 (before and after the partial

[5]Kyushu is divided into a few parts by mountains, and they have been relatively independent from each other. This is reflected in dialects. In fact, Kyushu has three dialects: "Hichiku", "Honichi" and "Satsugu" dialects (Tojo, 1927). The treatment group mainly belongs to the areas with "Hichiku" or "Satsugu" dialect, and the control group mainly belongs to the areas with "Honichi" dialect.

[6]In 1987, the national railway company in Japan was privatized and divided into seven private railway companies, of which the Kyushu Railway Company is one.

opening), and for 2008 and 2014 (before and after the entire opening). Note here that the project was planned and announced in 1973. However, "*the actual construction was subject to substantial timing uncertainty due to numerous budgetary and administrative delays, thus limiting the scope for anticipation effects*" (Bernard et al., 2019). Moreover, because the construction started in 1991, we consider that the effects of the project announcement and breaking ground, if any, had already been capitalized at the beginning of operation. Hence, our analysis is able to offer a bottom line of the overall effects of the Kagoshima route construction.

Construction of the Kagoshima route largely decreased the travel time between connected regions. The conventional wisdom of spatial economics holds that such economic integration will significantly affect agglomerations of economic activities (Fujita et al., 2001; Fujita and Thisse, 2013). As shown by Akamatsu et al. (2017), these effects crucially depend on the characteristics of the dispersion force. When the dispersion force is global as in the case of transporting goods and service to a distant, less-crowded market, integration accelerates concentration both across regions and within a region. When the dispersion force is local as in the case of urban congestion diseconomies, integration attenuates concentration both across regions and within a region.

Construction of the Kagoshima route undoubtedly fosters economic integration among regions within Kyushu, especially within its western part, and hence, would significantly affect metropolitan areas in Kyushu. Kyushu is a large island located in the western part of Japan. It has a population of 14.6 million, and has 17 metropolitan areas (MAs) defined by the Metropolitan Employment Areas (MEAs), which will be explained in detail below. Among its 17 MEAs, two are located in the Okinawa prefecture, which is a remote island prefecture, is much less likely to be affected by construction of the Kagoshima route. Hence, we exclude two MEAs from our analysis. Of the remaining 15 MEAs (Figure 2.2), six (the Fukuoka, Kurume, Omuta, Kumamoto, Yatsushiro, and Kagoshima MEAs) are located on the Kagoshima route. Accordingly, construction of the Kagoshima route is likely to have especially affected economic integration among these six MEAs. Within the six MEAs, the Fukuoka MEA, which is the largest MEA in Kyushu, is especially large, and the Kumamoto and Kagoshima MEAs are relatively larger than the remaining three MEAs. Hence, it is likely that the three small MEAs will fall into the agglomeration shadows of the three large MEAs to get shrunk in the dominance of the global dispersion force. In contrast, the three small MEAs will enlarge to make the six MEAs more equally sized in the dominance of the local dispersion force. Our analysis uncovers which scenario is relevant to agglomerations in Kyushu. We also explore the effects on smaller urban agglomerations defined by the Micropolitan Employment Areas (McEAs), which again will be explained below.

We investigate the impacts of construction of the Kagoshima route on the distribution of economic activities across urban agglomerations by examining changes in land prices. If a particular MEA can attract economic activities due to construction of the Kagoshima route, its land prices will rise. If it loses economic activities, its land prices will decline. The impacts of construction of the Kagoshima route capitalized in land prices are estimated using the hedonic approach a la Rosen (1974) via a difference-in-differences approach. Because the operation of the Kagoshima route occurred in two stages wherein the southern part of the line started operating in 2004 and the remaining part in 2011, we conduct the difference-in-differences for each step. In order to estimate the effects on land prices in MEAs, we set dummies that represent whether a point is located in each MEA. Moreover, in the estimation, we include characteristics of each point as control variables.

2.3 Data

Data on land prices comes from the Official Announcement of Land Prices published by the Ministry of Land, Infrastructure, Transport and Tourism (MILT). The Land Appraisal Committee of MILT selects a point (Officially, a standard site), asks two or more real estate appraisers to appraise these points, and judges and publicly announces the proper land price per square meter once a year.[7] It also describes each point's location, acreage, land-use zoning, building-area ratio, and floor-area ratio.[8] Note here that land prices used in this chapter have been appraised by real estate appraisers and are not transaction prices; although the latter can reflect specific factors about a point that are not explicitly captured by data, the former is standardized regarding such factors across neighborhoods.[9] Because our focus is not on a point's implicit specific factors, we use the former.[10] Kyushu contains eight prefectures (Fukuoka, Saga, Nagasaki, Kumamoto, Oita, Miyazaki, Kagoshima, and Okinawa). However, because Okinawa is a small isolated island prefecture located extremely distant from the other seven prefectures, we use land price data on other seven prefectures located on the main Kyushu island. Japan as a whole has approximately 26,000 points and the seven Kyushu prefectures have approximately 2,600 thousand points. We estimate the effects of the high-speed railway's partial opening by using data for the years 2001 and 2007, which are 3 years prior to and 3 years after the operation's start, respectively. Similarly, we estimate the effects of the entire opening by using data for the years 2008 and 2014. Each year, MILT replaces some of the points with new ones, and the share of common points between 2001 and 2007 is approximately 0.6 and that between 2008 and 2014 is approximately 0.3. The analysis data include 5,488 samples for the partial opening and 4,740 samples for the entire opening. Table 2.1 reports the descriptive statistics of the analysis sample.

[Table 2.1]

For the metropolitan areas, we used the Urban Employment Areas (UEAs), which are defined similarly to the Consolidated Metropolitan Statistical Areas and derived from Japanese municipalities (see Kanemoto and Tokuoka 2002). We use the UEAs based on the 2010 Population Census of Japan.[11] A UEA having the densely inhabited district (DID) population larger than 50,000 in the central city is called a Metropolitan Employment Area (MEA), and one having the DID population of 10,000-50,000 in the central city area is called a Micropolitan Employment Area (McEA). Figure 2.2 (resp. Figure 2.3) is a map of the MEAs (resp. McEAs) in Kyushu

[7] It appraises the points as of January 1 and announces prices in late March every year.

[8] Information on the last two are available if the point is located in the town planning area.

[9] Real estate appraisers take actual transaction cases into account for appraisal. Thus, prices of land where there are few transactions may have measurement errors. Even if we exclude points outside of city planning areas, where there are few transactions, from observations, results are mostly unaltered.

[10] Land transaction prices are available in the Land General Information System maintained by MILT.

[11] During the periods of our focus, we experienced a large number of municipality mergers. For our analysis, we converted the municipality information of the standard points into the one observed on October 1, 2010.

and the Kagoshima route of Kyushu Shinkansen.[12] Table 2.2 and Table 2.3 summarize the characteristics of MEAs/McEAs in Kyushu in terms of population, natural environment and economic base. Because detailed addresses for the points are available, we can identify whether a particular point is located within the areas of any UEA, and make a dummy that takes one if it is located within a particular UEA. The area of our focus has 15 MEAs and 22 McEAs. In the first step of constructing the Kagoshima route, the line passed two MEAs (Yatsushiro and Kagoshima) and two McEAs (Minamata and Satsumasendai) located on the line, whereas in the second step, it passed additional four MEAs (Fukuoka, Kurume, Omuta, and Kumamoto) and two McEAs (Tosu and Tamana).

[Figure 2.2, Figure 2.3, Table 2.2 and Table 2.3]

In order to identify the areas impacted by the construction of the Kagoshima route, we also use data on a point's distance to its nearest Shinkansen station on the Kagoshima route.[13] The MILT provides geographical data on the railroad network for the year 2015, which, combined with the address information of points, we used to compute the linear distance between each point and its nearest Shinkansen station on the Kagoshima route. For the partial opening of the Kagoshima route, we derived the distance to the nearest station for those from Shin-Yatsushiro to Kagoshima-Chuo Stations, and for the entire opening, we derived the distance to the nearest station for those from Hakata to Kagoshima-Chuo Stations.

2.4 Econometric Specification

This chapter uses the difference-in-differences approach to uncover the impacts of construction of the Kagoshima route on urban agglomerations. For each step of the route's construction, we estimate three specifications. The first specification examines the impacts on urban agglomerations, that have at least one Shinkansen station, as a whole. The second one investigates differences in the impacts among these agglomerations. The last one explores the impacts on each urban agglomeration more in detail. In all specifications in each step, we use data for 3 years before and 3 years after the operation's start. In all estimations, the dependent variable is $\ln(p_{ijt})$, which denotes the logarithm of land price of point i located in municipality j at time t. We denote the after-treatment dummy by w_t, the point i's attribute vector by X_{it}, the time-invariant unobserved characteristics of municipality j by c_j, the year dummy by p_t, and the idiosyncratic error term by u_{ijt}. Because the operation of the Kagoshima route occurred in two steps, w_t takes zero if $t = 2001$ and one if $t = 2007$ for the partial opening, and zero if $t = 2008$ and one if $t = 2014$ for the whole opening. X_{it} includes the distance from the nearest train station, building-area ratio, floor-area ratio, acreage, land-use zoning dummies (no regulation/residential purpose/commercial purpose/industrial purpose) and supply system dummies (water/gas/drain).[14]

[12] As one can see from the figure, the Kitakyusyu MEA locates at the gate from the main-land Japan. Moreover, when considering distance to the closest Shinkansen station, particular sites in the Saga and Nagasaki MEAs have a short linear distance although they have a much longer road distance because of the existence of the sea. In order to remove the effects of such special locational characteristics, we also conducted estimation without the Kitakyusyu, Saga, and/or Nagasaki MEAs and confirmed our results are unaltered. See Appendix C for details.

[13] In computing distances, we use the geodist command available in Stata.

[14] The distance from the nearest train station is calculated by linear distance.

In the first specification, which we call Analysis 1, we aim to estimate the impacts on urban agglomerations with the Shinkansen stations as a whole. We specify the estimation equation as

$$\ln(p_{ijt}) = \alpha + \beta_{mea} w_t \sum_{l=1}^{L} Z_{j,l} + \beta_{mcea} w_t \sum_{s=1}^{S} Z_{j,s} + X_{it}\gamma + c_j + p_t + u_{ijt}, \tag{2.1}$$

where l and s represent the MEA and McEA that have at least one Shinkansen station, respectively, and $Z_{j,l}$ (resp., $Z_{j,s}$) is the dummy that takes one if municipality j is included in MEA l (resp., McEA s) and zero otherwise. Hence, $\sum_{l=1}^{L} Z_{j,l}$ (resp., $\sum_{s=1}^{S} Z_{j,s}$) takes one if municipality j is included in any MEA (resp., McEA). Our DID estimators are β_{mea} and β_{mcea}: if the estimated β_{mea} (resp., β_{mcea}) is positive, construction of the Kagoshima route has boosted land prices in MEAs (resp., McEAs) constructed by the new Shinkansen.

Note here that identification of the DID estimator requires no correlation between the independent variable related to the DID estimator and the error term. However, as explained by Ministry of Land, Infrastructure and Tourism, profitability and investment efficiency are required for the construction of the Shinkansen, implying that the location choice of the Shinkansen might possibly relate to time-invariant unobserved characteristics. In order to control for such a correlation, we included a fixed effect for each municipality, c_j, in the estimation equation.

In the second specification, which we call Analysis 2, we extend the first specification to examine differences in the impacts among urban agglomerations that have at least one Shinkansen station. We specify the estimation equation as

$$\ln(p_{ijt}) = \alpha + \sum_{l=1}^{L} \beta_l w_t Z_{j,l} + \sum_{s=1}^{S} \beta_s w_t Z_{j,s} + X_{it}\gamma + c_j + p_t + u_{ijt}, \tag{2.2}$$

where l and s represent the MEA and McEA that have at least one Shinkansen station, respectively. $Z_{j,l}$ (resp., $Z_{j,s}$) is the dummy that takes one if municipality j is included in MEA l (resp., McEA s) and zero otherwise. Note here that we estimate the DID estimators β_l and β_s for each MEA and McEA, respectively. This reflects our idea that the impacts might differ depending on the size and location of the MEA/McEA along the route.

In the last specification, which we call Analysis 3, we look in greater detail at the impacts on each urban agglomeration by using data on linear distance between the point and the nearest Shinkansen station on the Kagoshima route. For this purpose, we categorize the points into six groups according to their distance to the nearest Shinkansen station. Each category has a range of 5km, i.e., the first category consists of points with a distance shorter than 5km, the second category consists of those with the distance equal to or greater than 5km and less than 10km, and so on, such that the final (6th) category consists of those with a distance equal to or greater than 25km. Letting $D_{d,i}$ denote the dummy that takes one if point i is included in category d and zero otherwise, we specify the estimation equation as

$$\ln(p_{ijt}) = \alpha + \sum_{l=1}^{L}\sum_{d=1}^{6} \beta_{d,l} w_t D_{d,i} Z_{j,l} + \sum_{s=1}^{S}\sum_{d=1}^{6} \beta_{d,s} w_t D_{d,i} Z_{j,s} + X_{it}\gamma + c_j + p_t + u_{ijt}. \tag{2.3}$$

where l and s represent the MEA and McEA that have at least one Shinkansen station. Again, $Z_{j,l}$ (resp., $Z_{j,s}$) is the dummy that takes one if municipality j is included in MEA l (resp., McEA s) and zero otherwise. Because the DID estimator $\beta_{d,l}$ and $\beta_{d,s}$ now depend on the distance to the nearest Shinkansen station, this specification allows us to uncover the different impacts on locations within each MEA/McEA with the Shinkansen stations.

2.5 Estimation Results

2.5.1 Analysis 1

Table 2.4 presents the estimation results of the first specification (2.1).

[Table 2.4]

In this analysis, we examine the impacts on urban agglomerations with the Shinkansen stations as a whole by estimating β_{mea} and β_{mcea} for each step of construction of the Kagoshima route. Columns (1)-(3) show the results for the partial route's opening in 2004 and columns (4)-(6) show the results for the entire opening in 2011. In columns (1) and (4) (resp., columns (2) and (5)), we include only the DID estimator for the MEAs (resp., the McEAs) that have at least one station along the Kagoshima route whereas columns (3) and (6) include the DID estimator for both. The signs of the estimated coefficients for control variables are as expected, and we omit them from Table 2.4 because they do not fall within our focus.[15][16]

Partial Opening of the Kagoshima route

In columns (1) and (3), the DID estimator for the MEAs with the Shinkansen stations (β_{mea}) is significantly positive. For the McEAs, the DID estimator (β_{mcea}) is also positive but insignificant both in columns (2) and (3). The partial opening of the Kagoshima route increased land prices of the MEAs with the Shinkansen stations by an average of around 11.3% whereas it affected that of the McEAs only insignificantly. These results imply that the economic activities were attracted to MEAs having stations along the Kagoshima route.

Opening of the entire Kagoshima route

Similarity to the partial opening, the DID estimators for the MEAs having the stations at the opening of the entire route are positive and significant, as shown in columns (4) and (6). The opening of the whole Kagoshima route raised land prices in the MEAs with the Shinkansen stations in Kyushu on average by around 8.1%. Regarding the McEAs, the estimated coefficient is negative, but again insignificant, as shown in column (5) and (6). In short, the opening of the entire route in 2011 on average benefited large urban agglomerations on the route.

[15] The results for control variables are provided in Appendix A.

[16] In our analysis, we cluster the standard errors at a municipality level. Even if we cluster the standard errors at an urban employment area level, results are mostly unaltered. We have only 38 clusters when clustering at an urban employment area level. Because this figure is slightly smaller than the threshold number of clusters (i.e., 42) required for clustering standard errors given by Angrist and Pischke (2008), we report the results under clustering at a municipality level.

In the analysis of the opening of the entire route, one concern is the effect of the global financial crisis in 2008. Because it is included in our analysis period, the estimation results may include the crisis's effect on land prices. In order to eliminate this effect, we replace data for the year 2008 with data for the year 2009 in columns (7). Although the absolute values of the estimated coefficients for MEAs become slightly smaller compared with those in columns (6), the significance of the coefficients does not change, which implies that the global financial crisis was not responsible for the results for the opening of the whole route.

2.5.2 Analysis 2

Results of Analysis 1 show that the MEAs on the Kagoshima route attracted economic activities as a whole through the opening and extension. However, the impacts on each MEA/McEA can be heterogeneous as the literature of New Economic Geography shows. In the second analysis, we try to exploit such differences more in detail. Table 2.5 shows the estimation results of the second specification (2.2).

[Table 2.5]

The DID estimators are β_l, and β_s, which are the DID estimator for each MEA/McEA having stations on the Kagoshima route. Columns (1)-(3) show the results for the partial opening in 2004 and columns (4)-(6) show the results for the entire route's opening in 2011. Columns (1) and (4) (resp., columns (2) and (5)) include only the DID estimators for the MEAs (resp., those for the McEAs), and columns (3) and (6) includes both.[17]

Partial opening of the Kagoshima route

After the partial opening, the Shinkansen passed through two MEAs (Kagoshima and Yatsushiro) and two McEAs (Minamata and Satsumasendai). Columns (1) and (3) show the DID estimator for MEAs having stations on the Kagoshima route. The estimated coefficient for the Kagoshima MEA is positively significant and its value is 0.132, implying that land prices in the Kagoshima MEA increased by 13.2%. On the contrary, the estimated coefficient for the Yatsushiro MEA is negatively significant, and its value is around -0.079, implying that land prices in this MEA decreased by 7.9%.

The results for the McEAs having stations on the Kagoshima route are shown in columns (2) and (3). The estimated coefficient for the Minamata McEA is negative and significant. As seen in column (3), the Minamata McEA experienced changes in land price by -5.4%, which is slightly smaller in magnitude than the changes in the neighboring MEA, i.e., Yatsushiro MEA. The estimated coefficient for Satsumasendai McEA is weakly significant. If we take the figure in column (3), the McEA experienced land price rises of 5.4%. In total, the partial opening attracted economic activities to the Kagoshima MEA from other small MEAs (especially from Yatsushiro MEA and its neighboring McEA). Because the Kagoshima MEA is the largest urban agglomeration among those on section of the Kagoshima route operative in 2004, the partial

[17] Again, the results for control variables are as expected and are provided in Appendix A.

opening of the Kagoshima route expanded the already-large urban agglomeration at the expense of other smaller urban agglomerations on the route.

Opening of the entire Kagoshima route

After the opening of the entire route, the Shinkansen traveled through four other MEAs (Fukuoka, Kurume, Omuta, and Kumamoto) and two McEAs (Tosu and Tamana), each of which has at least one station on the route. Columns (4) and (6) show the DID estimator for MEAs. The estimated coefficients of the DID estimators are positive and significant for the larger MEAs (Fukuoka and Kumamoto). The Fukuoka MEA, the largest MEA, had the largest coefficient of 0.107, indicating that land prices rose by 10.7%. The Kumamoto MEA, the second largest MEA, followed the Fukuoka MEA with a coefficient of 0.085, indicating a 8.5% increase in land prices. The estimated coefficients of the DID estimators are negative and significant for smaller MEAs (Omuta and Yatsushiro). The Yatsushiro MEA has a coefficient of -0.067, implying that land prices in this MEA decreased by 6.7%. The estimated coefficient for the Omuta MEA is -0.057, implying that land prices in this MEA decreased by 5.7%, but it is not very significant. The estimated coefficients for the intermediate-size MEAs (Kurume and Kagoshima MEAs) are significantly positive. Those MEAs saw increases in land prices by 5.7% and 7.0%, respectively. The DID estimator for McEAs are provided in columns (5) and (6), where it can be seen that the estimated coefficients for the McEAs are not very significant except Minamata McEA.

Thus, our results induce us to support the conclusions that the opening of the entire route attracted economic activities to large urban agglomerations (i.e., MEAs) on average, and drastically changed the distribution of economic activities among the MEAs having stations along the route: larger MEAs gained and smaller MEAs lost. In contrast, it affected small urban agglomerations (i.e., McEAs) only insignificantly. Note however, we must examine the effects on different locations within each urban agglomeration in order to know the integration accelerated/attenuated agglomeration. Hence, we next divide the areas within each city into several categories by distance from the Shinkansen station and estimate the effects by distance categories.

In order to eliminate the effect of the financial crisis, we again replace data for the year 2008 with data for the year 2009 in column (7) of Table 2.5. Although the absolute value of the estimated coefficients for MEAs has slightly changed compared with those in column (6), the significance of the coefficients does not change, which implies that the global financial crisis is not responsible for the results for the whole route's opening.

2.5.3 Analysis 3

In the third analysis, we examine the effects on each urban agglomerations more in detail by focusing on the distance of each point from the nearest Shinkansen station. Tables 2.6-2.8 provides the estimation results of the third specification (2.3).

[Tables 2.6, 2.7 and 2.8]

Here, the DID estimators are $\beta_{d,l}$ and $\beta_{d,s}$, where the subscript d represents the category for distance from the nearest Shinkansen station. We set six distance categories from the nearest

Shinkansen station: 0-5km, 5-10km, 10-15km, 15-20km, 20-25km, and over 25km. Tables 2.6 and 2.7 show the estimated coefficients for the partial opening and entire opening, respectively.

Partial opening of the Kagoshima route

Table 2.6 shows the estimation results of Analysis 3 for the partial opening. For the Yatsushiro MEA, the coefficients for 0-5km and 5-10km distance categories are negative and significant, taking a value of -0.085 and -0.130 respectively. Hence, areas closest to the Shinkansen station (0-5km) saw land prices drop by around 8.5%. For the Minamata McEA, which is close to the Yatsushiro MEA, a distance category exists only for 0-5km and its estimated coefficient is -0.052; this is significant, implying decreases in land prices by around 5.2%.

Inside the Kagoshima MEA, only the 0-5km and 15-20km areas saw land prices increase; other areas experienced decreases. Although the Kagoshima MEA as a whole has attracted economic activities, areas within the MEA that have attracted such activities are limited, and they are relatively close to the Shinkansen station. For areas distant from a Shinkansen station by more than 25km, land prices have declined dramatically. For the Satsumasendai McEA, which is close to the Kagoshima MEA, land prices increased only in the 5-10km and 15-20km category. Similar to the Kagoshima MEA, areas that attracted economics activities inside the McEA are limited. Because the neighborhood of the Shinkansen station in these MEA and McEA originally had higher land prices than other areas, we can conclude that the partial opening accelerated agglomeration within each of these cities.[18] Thus, our results here is in line with the view of the global dispersion force a la Akamatsu et al. (2017).

Opening of the entire Kagoshima route

Table 2.7 presents the estimation results of Analysis 3 for the opening of the whole route. We can see strong positive and significant effects on land prices for the 0-5km category in the Fukuoka, Kumamoto, and Kagoshima MEAs. The estimated values for the 0-5km category are 0.312 for the Fukuoka MEA, 0.162 for the Kumamoto MEA, and 0.312 for the Kagoshima MEA, resulting in land price increases of 31.2%, 16.2%, and 31.2%, respectively. However, such positive effects become weaker with distance to a Shinkansen station, and in the Kumamoto and Kagoshima MEAs, the effects become even negative with high significance for areas distant from a Shinkansen station. Because the neighborhood of the Shinkansen station in these MEAs originally had higher land prices than other areas, we can conclude that the full opening accelerated agglomeration within each of these MEAs. Thus, our results is again in line with the view of the global dispersion force a la Akamatsu et al. (2017).

These positive effects propagate to the MEAs and McEAs close to the Fukuoka, Kumamoto, and Kagoshima MEAs (i.e., Omuta and Kurume MEAs and Tosu McEA, which are close to the Fukuoka MEA, Tamana McEA, which is close to the Kumamoto MEA, and Satsumasendai McEA, which is close to the Kagoshima MEA), and the effects are positive and significant for areas close to the Shinkansen stations in these urban agglomerations. However, again, the positive effects are limited to areas close to Shinkansen stations, and they turn to be negative

[18] See Appendix B for figure showing the relationship between land price and distance from the Shinkansen station for relevant cities.

or insignificant as distance to a Shinkansen station increases. In contrast, the Yatsushiro MEA experienced large decreases (-33.9%) in land prices for the 5-10km category.

We can summarize our findings as follows. First, the opening of the entire Kagoshima route attracted economic activities to already large agglomerations with the Shinkansen stations along the Kagoshima route. Secondly, it also accelerated agglomeration within each large agglomeration.

In order to eliminate the effect of the financial crisis, we again replace data for the year 2008 with data for the year 2009 in Table 2.8. We obtain very similar results to those shown in Table 5, which implies that the global financial crisis is not responsible for the results from the opening of the entire route.

2.6 Econometric Issues

2.6.1 Placebo Test

In order to check if the treatment and control groups share common time trends in land prices before the construction of Shinkansen, we implement placebo tests. For the placebo test regarding the partial opening in 2004, we use data for the years 1994 and 2000 as if the partial opening had occurred at the beginning of 1997. Similarly, for the test regarding the entire opening in 2011, we use data for the years 2002 and 2008 as if the entire opening had occurred at the beginning of 2005. The following results imply that the common trend assumption cannot be rejected.

Table 2.9 shows the results of estimating the same specification as those shown in columns (1) - (6) in Table 2.4 (i.e., Analysis 1).

[Table 2.9]

From Table 2.9, we know that the DID estimators of the MEAs and the McEAs with the Shinkansen stations are insignificant for both the partial opening and the entire opening. An exception is column (2) which includes only the DID estimator for the McEAs with the stations but it is also weakly significant. Hence, we cannot reject the null hypothesis that the treatment and control groups have same time trends in land prices.

The common trend assumption can be verified also by Figure 2.4, which shows the average land prices in treatment and control groups for each event. Here, the treatment group at the partial opening includes points in the two MEAs (Yatsushiro and Kagoshima) and the two McEAs (Minamata and Satsumasendai), while the treatment group at the entire opening includes points in the six MEAs (Fukuoka, Omuta, Kurume, Kumamoto, Yatsushiro and Kagoshima) and the four McEAs (Tosu, Minamata, Tamana, Satsumasendai). Each control group includes remaining points beside points in the treatment group for each event. The red vertical line shows the year of the partial opening (2004) or the entire opening (2011). Before each event, both treatment and control groups have declining trends in land prices, and they moved in parallel. They coincide with the declining trends in the average land price in Japan. After the partial opening, the decrease trend in average land price of the treatment group is slower than the control group. After the entire opening, only the treatment group experienced increases in land prices.

[Figure 2.4]

2.6.2 Sub-sample analysis by land-use zoning

Japan uses land-use zoning as a part of its urban land-use planning system, which is implemented by local municipalities under the City Planning Law. Twelve categories of land use zones are defined, and each of them specifies the use of buildings that are allowed to be constructed in a zone. These categories can be generally divided into three groups: residential, commercial, and industrial uses. In this subsection, we construct three sub-samples, each of which corresponds to one group, and conduct analyses with specifications (2.1) and (2.2) for each sub-sample.[19][20] This enables us to study which types of location have been more intensively affected by construction of the high-speed railway.

Table 2.10 shows the estimation results of the first specification (2.1). Here, we estimate the DID estimators of MEA/McEA having stations on the Kagoshima route. Columns (1)-(4) show the results for the partial opening in 2004 and columns (5)-(8) show the results for the opening of the entire route in 2011. Columns (1) and (5) are the baseline results that come from columns (3) and (6) in Table 2.4. Columns (2), (3), and (4) (and columns (6), (7), and (8)) show the results of sub-sample analyses for residential, commercial, and industrial uses, respectively.

[Table 2.10]

For the partial opening in 2004, the DID estimators are significantly positive only for the residential use; surprisingly they are insignificant for the commercial use as well as for the industrial use. This findings implies that the partial opening positively affected land prices in residential areas but not in commercial and industrial areas. On the other hand, the opening of the entire route in 2011 increased land prices in residential and commercial areas and insignificantly affect industrial area. The estimated increases are similar between these two types of land uses, taking values of around 7.0%. Because Shinkansen is not designed for freighting, it would be reasonable that the construction of Shinkansen does not affect industrial areas. In contrast, we expect that it would significantly affect residential and commercial areas.

Then, why did the partial opening have an insignificant effect on commercial areas insignificantly? We raise two possibilities. First, it is possible that land prices in commercial areas had responded earlier than the actual opening. The historical trends in Japanese land prices show that land prices in commercial areas have responded to events such as the "bubble economy" more quickly than land prices in residential areas (Sato, 2014). Because people expected the opening of the Kagoshima route, land prices in commercial areas might have increased prior to the opening in 2004. Thus, the estimated impact shown in column (1) in Table 2.10 may be biased downward compared with the actual impact. Second, the effects of firm sorting might be responsible for the insignificance. The partial opening could have intensified competition among

[19] Although each specification shown in Section 5 includes land-use zoning dummies, they are omitted in this section.

[20] Appendix D conducts more detail sub-sample analysis.

firms, as shown by Melitz (2003), forcing firms with lower productivity to exit the market and enabling firms with higher productivity to make higher profits. Such a sorting effect might be responsible for the unclear effect on commercial land prices. In contrast, such competition does not exist in residential areas, resulting in significantly positive DID estimators. This possibility is consistent with the findings of Bernard et al. (2019) that the partial opening of the Kagoshima route increased firms' sales prices.

Table 2.11 presents the results of sub-sample analyses under the second specification (2.2). Again, columns (2)-(4) (resp., columns (6)-(8)) show the results for the partial opening (resp. the entire opening) for residential, commercial, and industrial areas, respectively. Columns (1) and (5) are the baseline results that come from columns (3) and (6) in Table 2.5.

[Table 2.11]

For the partial opening in 2004, signs on the estimated coefficients are almost the same across columns (1)-(3). However, significant increases in land prices were observed only in residential areas of the Kagoshima MEA and Satsumasendai McEA. In contrast, the significant decreases in land prices were observed only in commercial areas of the Yatsushiro MEA. These results imply that the partial opening induced residents to move toward larger MEAs/McEAs and firms to leave smaller MEAs. For the entire route's opening in 2011, although signs of the DID estimators are again almost the same across columns (5)-(8), their significance varies. For residential areas (column (6)), only MEAs having stations on the Kagoshima route experienced significant impacts, which were positive for relatively larger MEAs such as Fukuoka and Kumamoto MEAs, and negative for relatively smaller MEAs such as Omuta and Yatsushiro MEAs. These results imply that the opening of the entire route induced residents to move from smaller to larger agglomerations. For commercial areas (column (7)), three largest MEAs, i.e., Fukuoka, Kumamoto, and Kagoshima MEAs, and McEAs close to these MEAs such as Tosu McEAs experienced significant positive impacts, implying that firms are attracted to larger MEAs and their neighborhoods by the opening of the entire route.

2.6.3 Dynamic effects

It is possible that the impacts of high-speed rail construction emerge gradually. To capture such a process, we conduct the difference-in-differences estimation using multiple treatment periods. We use land price data for 5 years before and after each event: we use data from 2000 to 2009 for the partial opening and data from 2007 to 2016 for the entire opening. We modify the first specification (2.1) as

$$\ln(p_{ijt}) = \alpha + \sum_{t=1}^{T} \beta_{mea,t} Y_t \sum_{l=1}^{L} Z_{j,l} + \sum_{t=1}^{T} \beta_{mcea,t} Y_t \sum_{s=1}^{S} Z_{j,s} + X_{it}\gamma + c_j + p_t + u_{ijt},$$

where $t \in \{0, 1, ..., T\}$ represents year. DID estimators ($\beta_{mea,t}$ and $\beta_{mcea,t}$) are the coefficients of year dummies multiplied by the MEAs/McEAs with the Shinkansen stations dummy. Y_t takes one if the year is t. We exclude the first year of each data from the year dummies to avoid collinearity. Thus, we estimate nine $\beta_{mea,t}$ and nine $\beta_{mcea,t}$ for each event, where the years 2000 and 2007 are the base years for the partial opening and the entire opening, respectively. Note here that land prices are evaluated on January 1st of each year in our data, and the day of the partial opening was March 13, 2004, and that of entire opening was March 12, 2011 Hence, $\beta_{mea,2004}$, $\beta_{mcea,2004}$, $\beta_{mea,2011}$ and $\beta_{mcea,2011}$ show the effects on land prices before the event. Each event happened between (base year + 4 year) and (base year + 5 year).

Table 2.12 presents the results. For the partial opening in 2004, the DID estimators of MEAs are significantly positive throughout the period. The estimator values are small before the opening. This implies that land prices in MEAs increased compared with prices in 2000 even before the partial opening. This could be due to early responses of commercial land prices as we discussed in Section 6.2. After the partial opening, the values of estimators gradually increased. This result also implies that the effect of the partial opening was not a temporary shock but a persistent one. In contrast, the DID estimators of McEAs are significant only for 2003 (i.e., base year + 4 year). However, we observe no significant effects on land prices in McEAs after the event. For the opening of the entire route in 2011, the DID estimators of MEAs before the event are only weakly significant. This result shows that the common trend assumption is highly satisfied for the opening of the entire route. After the event, they are significantly positive and the impacts continue to increase until the end of the period. We observe insignificant effects on McEAs throughout the periods.

[Table 2.12]

2.6.4 Equilibrium effects

The opening and extension of the Kagoshima route induced people and firms to move to agglomerations containing a Shinkansen stations from those without a station. This increased land demand and land prices in the former, implying that after the event, the positive slope of the hedonic curve on land price becomes steeper. Kuminoff and Pope (2014) showed that such an equilibrium effect can be a source of bias in DID estimation. In our analyses, we used the specifications wherein the marginal willingness to pay for land in agglomerations with Shinkansen stations is constant over time. We here follow Kawaguchi and Yukutake (2017) in considering the equilibrium effects and alter the first specification (2.1) as

$$\ln(p_{ij0}) = \alpha + \beta_{mea,0}(\text{MEA}_0) + \beta_{mcea,0}(\text{McEA}_0) + X_{i0}\gamma + c_j + p_0 + u_{ij0},$$
$$\ln(p_{ij1}) = \alpha + \beta_{mea,1}(\text{MEA}_1) + \beta_{mcea,1}(\text{McEA}_1) + X_{i1}\gamma + c_j + p_1 + u_{ij1},$$

where the subscript 0 represents the period before the opening and subscript 1 represents the period after the opening. "MEA" and "McEA" are dummies for MEAs and McEAs with Shinkansen stations, respectively, implying that $\beta_{mea,t}$ and $\beta_{mcea,t}$ represent the marginal willingness to pay for land in MEAs and McEAs with Shinkansen stations at time t ($t = 0, 1$), respectively. Note here that we assume that the opening and extension did not change the unit price of land characteristics, keeping γ constant over the periods. If we take the difference of these equations, we obtain the first-difference equation as

$$\begin{aligned}\Delta ln(p_{ij}) &= \beta_{mea,1}\Delta\text{MEA} + (\beta_{mea,1} - \beta_{mea,0})\text{MEA}_0 \\ &\quad + \beta_{mcea,1}\Delta\text{McEA} + (\beta_{mcea,1} - \beta_{mcea,0})\text{McEA}_0 \\ &\quad + \Delta X_i\gamma + \delta + \Delta u_{ij},\end{aligned} \tag{2.4}$$

where Δ is the difference operator and $\delta = p_1 - p_0$. As shown by Kuminoff and Pope (2014), if we omit MEA_0 (resp., McEA_0) from independent variables, we face conflation bias, i.e., $\beta_{mea,1}$ (resp., $\beta_{mcea,1}$) estimates the mixture of the effect of the event on land prices and the effect of changes in the hedonic curve. Hence, in general, the capitalization effects estimated in the DID analysis are not equal to the marginal willingness to pay. However, MEA_0 and McEA_0 are conceptually zero for all points in our analysis because no stations are available prior to the event. Hence, we can conclude that the equations to be estimated do not change.

2.6.5 Train types

After the opening of the entire route, three types of trains run on the Kagoshima route: "Mizuho," "Sakura," and "Tsubame" (Table 2.13). They differ in terms of number of stations where they stop, travel time, and service frequency.[21] The fastest train, "Mizuho," stops only at Hakata, Kumamoto and Kagoshima-Chuo stations, which are located in three largest MEAs in Kyushu. The service frequency is around eight times per day along both the in-bound and out-bound lines, and it runs mainly in the morning and the evening. "Sakura" is the second fastest and most frequent service. It stops at about half of the stations and runs once or twice every hour. "Tsubame" is the slowest service and stops at all stations. Some of them only run at the north part of Kagoshima route (i.e., between Kumamoto and Hakata stations), but here we focus on Tsubame, which runs along the entire route.

[Table 2.13]

It would be possible that the effects of high-speed railways on agglomerations depend on the types of trains that stop in them. In order to address this point, we conduct an additional analysis using the first specification with new two dummies: three types dummy, T_3, and two types dummy, T_2. T_3 takes one if the point is located in a MEA in which all types of trains stop (i.e., the Fukuoka, Kumamoto, and Kagoshima MEAs). T_2 takes one if the point is located in a MEA/McEA containing "Mizuho" and all "Sakura" stops.[22]

Interaction terms between these dummies and main independent variables are added to the first specification (2.1):

[21] Before the entire route's opening, the Kagoshima-route had only "Tsubame", which stops at every station on the route (West Japan Railway Company).

[22] Some "Sakura" stop at additional stations (see Table 2.13).

$$
\begin{aligned}
\ln(p_{ijt}) = \alpha &+ \beta_{mea} w_t \sum_{l=1}^{L} Z_{j,l} + \beta_{mea,T2} w_t T_2 \sum_{l=1}^{L} Z_{j,l} + \beta_{mea,T3} w_t T_3 \sum_{l=1}^{L} Z_{j,l} \\
&+ \beta_{mcea} w_t \sum_{s=1}^{S} Z_{j,s} + \beta_{mcea,T2} w_t T_2 \sum_{s=1}^{S} Z_{j,s} \\
&+ X_{it}\gamma + c_j + p_t + u_{ijt}.
\end{aligned}
$$

Note here that $\sum_{l=1}^{L} Z_{j,l}$ and $\sum_{s=1}^{S} Z_{j,s}$ are MEA and McEA with the Shinkansen stations dummies, respectively. Hence, the estimated values of β_{mea}, $\beta_{mea,T2}$, and $\beta_{mea,T3}$ (resp. β_{mcea} and $\beta_{mcea,T2}$) show the effect of the number of train types stopping in MEAs (resp. McEAs) on the MEAs' (resp. McEAs') land prices. Because no McEA has a station where all three types of trains stop, the above specification does not include an interaction term among three types dummy and McEA dummy.

Table 2.14 presents the result. Column (1) is baseline results that come from column (3) of Table 2.4. Column (2) includes the interaction terms with train types dummies. These columns show that the more types of trains that stop in a MEA is, the larger the increases in the MEA's land prices, implying positive effects upon land prices from the number of types of trains that stop. However, an interaction term between two types dummy and McEA dummy does not significantly affect land price. This implies that the existence of a Shinkansen station has no impact on McEAs' land prices, regardless of the number of types of trains stopping in a McEA.[23]

[Table 2.14]

2.7 Concluding Remarks

The opening and extension of a high-speed railway (Kagoshima route of Kyushu Shinkansen specifically) have connected several urban agglomerations in western Kyushu in Japan and have accelerated integration among them. We aimed to examine their effect on the distribution of economic activities across agglomerations by focusing on changes in land prices. For this purpose, we estimated hedonic price equations and conducted difference-in-difference analyses. Our results showed that the opening and extensions have increased land prices greatly in large metropolitan areas: thanks to the opening of the entire Kagoshima route, the largest metropolitan area in Kyushu experienced land price increases of around 10.7%, and the second largest metropolitan area saw prices rise by around 8.5%. However, if we extended the estimation by categorizing locations into different groups according to their distance from the nearest Shinkansen station, such positive effects are limited to areas close to Shinkansen stations. In contrast, small metropolitan areas located between large metropolitan areas experienced land price declines, indicating that they became stuck in the agglomeration shadows of the large

[23]Note here that some of the trains on the Kagoshima route run only between Hakata and Kumamoto stations. Hence, we also added a dummy representing that a station is located between the two stations. However, the estimated coefficient was not significant.

metropolitan areas. Thus, construction of the high-speed railway has accelerated concentration of economic activities in already-large urban agglomerations in Kyushu.

One important direction of future research would be to disentangle the mechanism behind the differences in impacts of high-speed rail by land-use zoning. Our results showed that residential and commercial zones experienced somewhat different impacts, whereas industry zones experienced non-significant impacts. Although we provided intuitive discussions, fully investigating such a mechanism would require us to build a multi-city model with networks among cities and conduct a structural estimation. This being of course beyond the scope of our current studies would be worth investigating in the future.

Figure 2.1: Kagoshima route and stations of the Kyushu Shinkansen

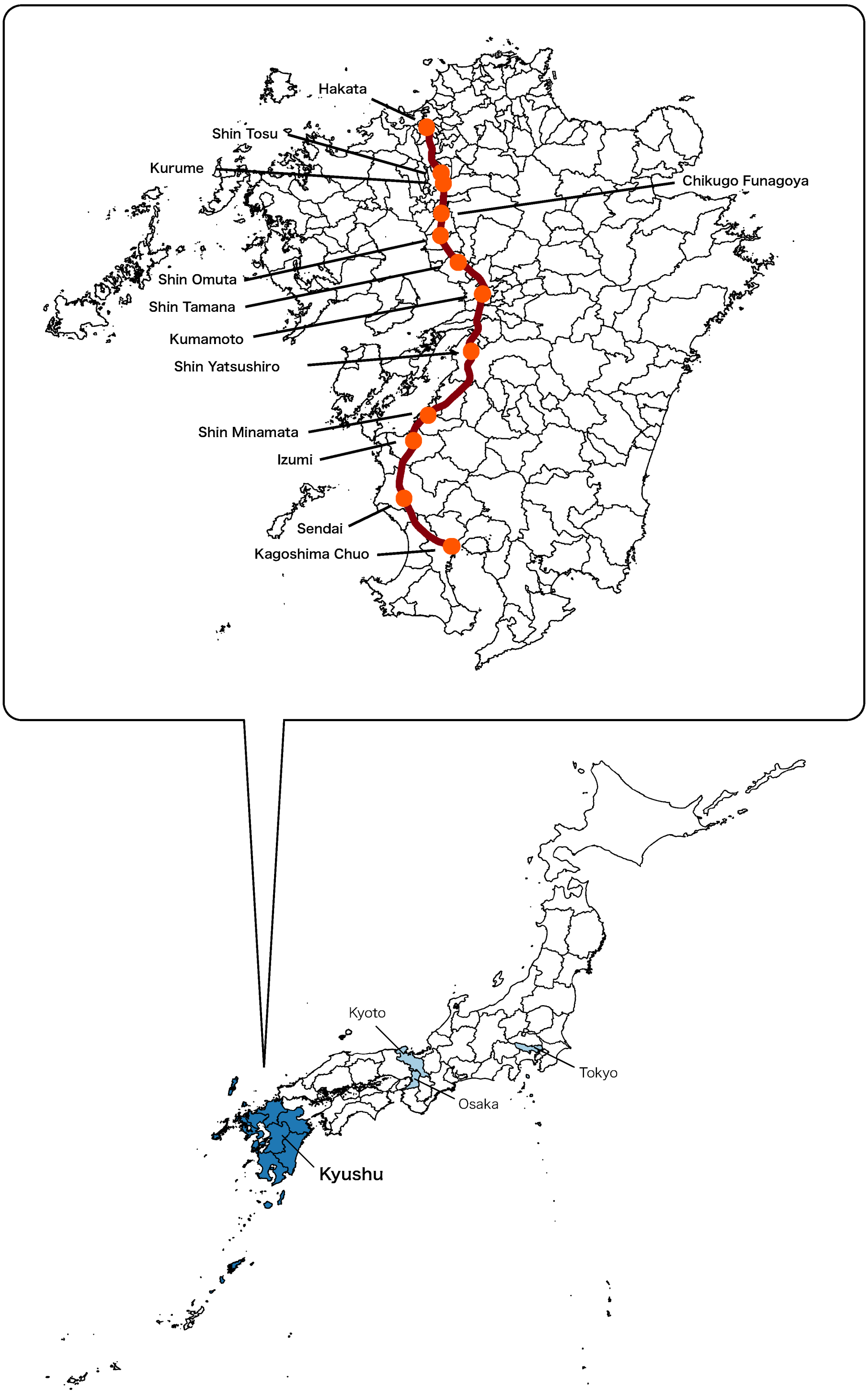


Figure 2.2: MEAs and the Kagoshima route of the Kyushu Shinkansen

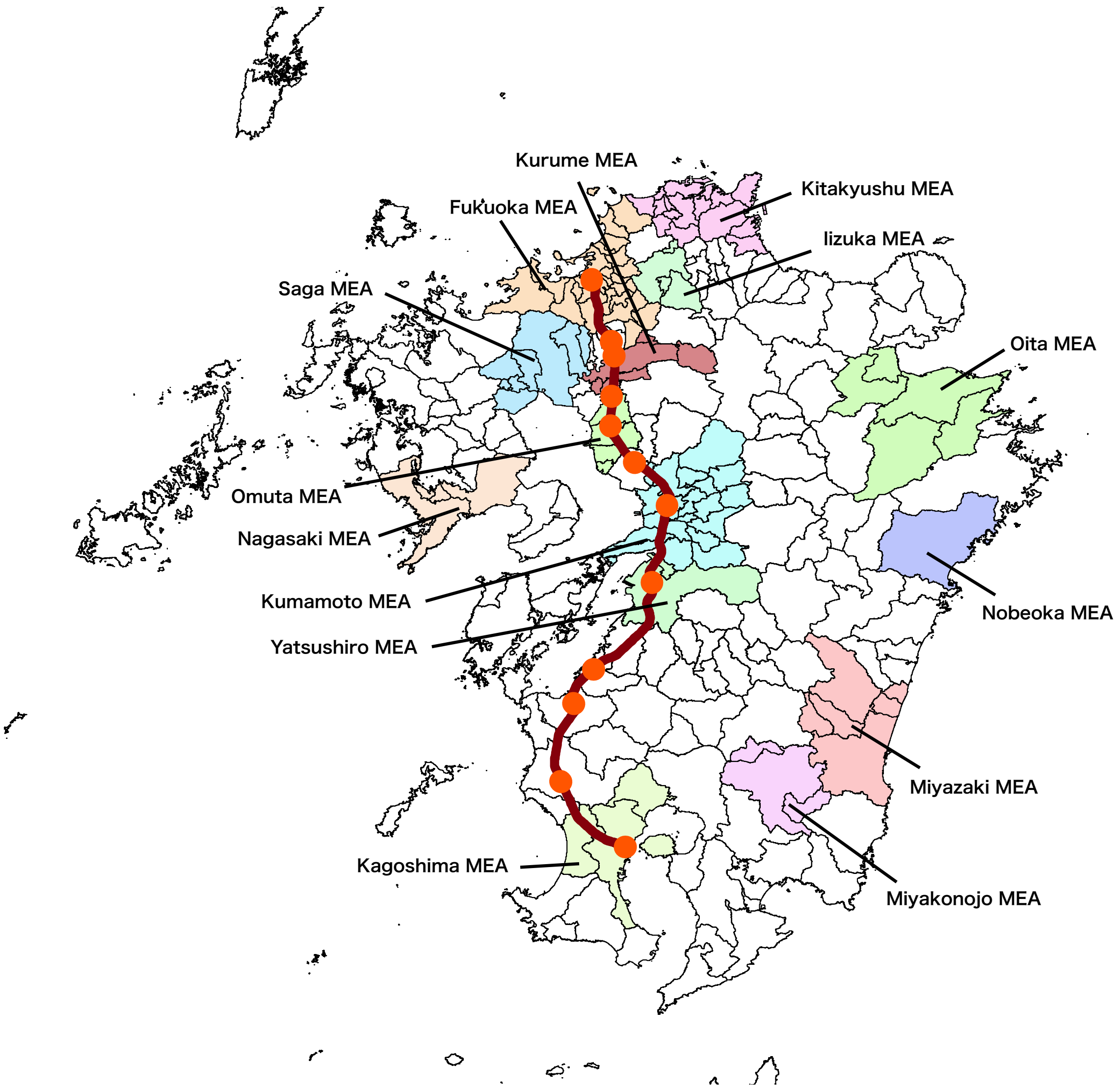


Figure 2.3: McEAs and the Kagoshima route of the Kyushu Shinkansen

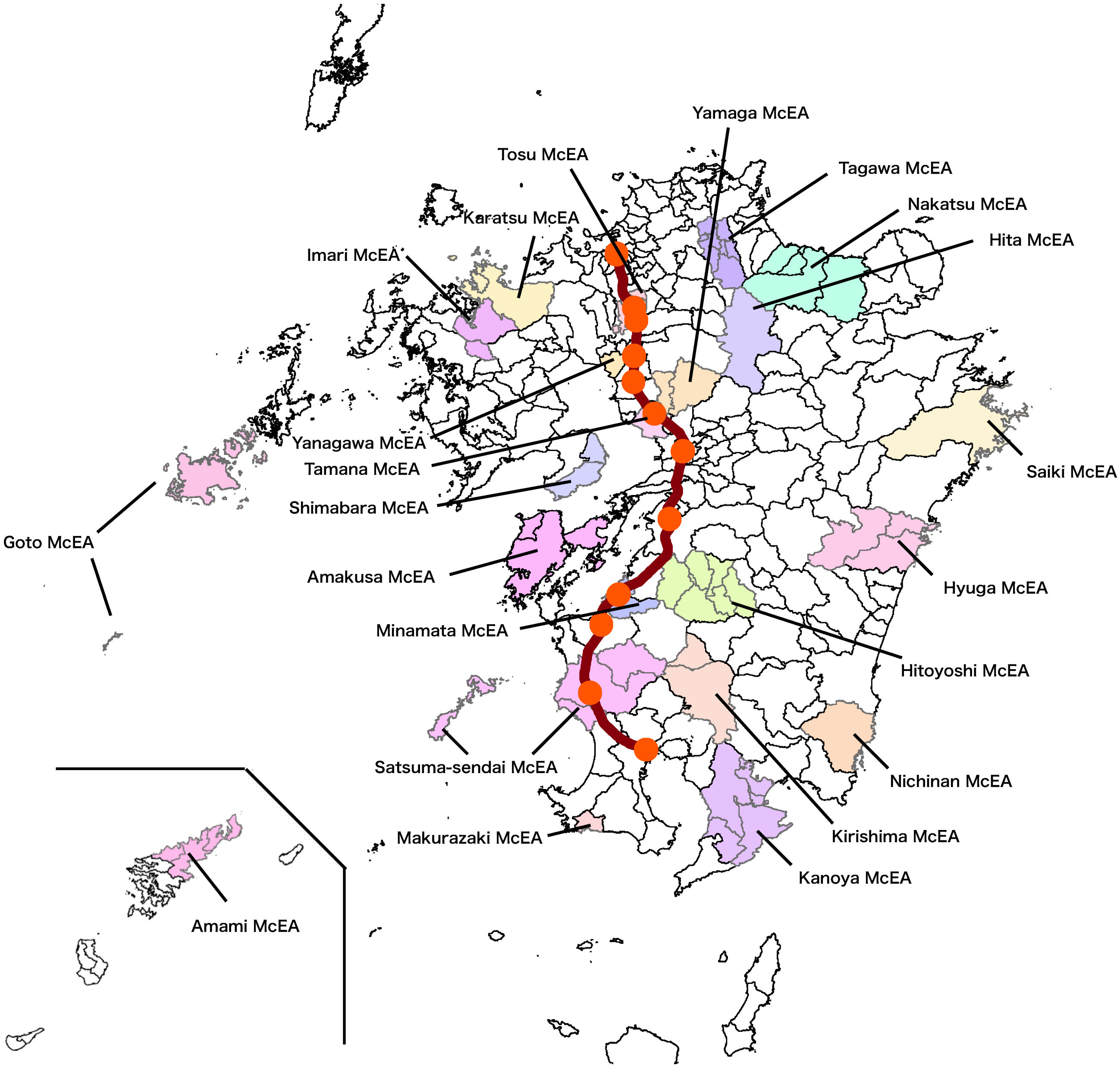


Figure 2.4: Land price trends

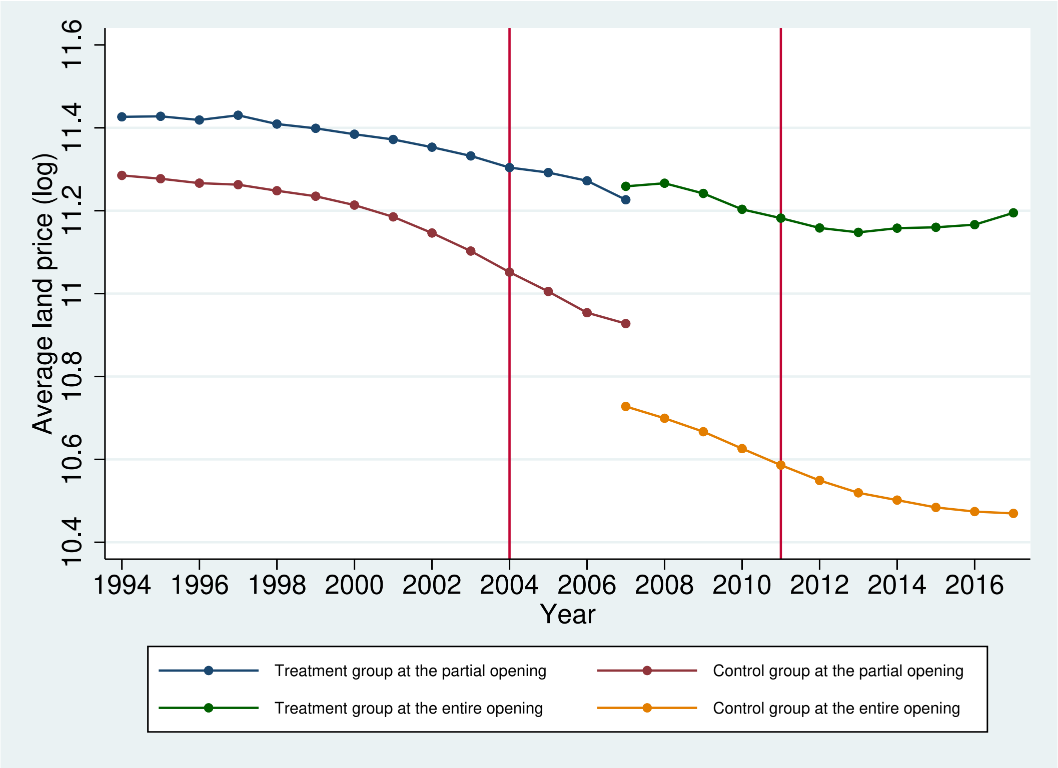


Table 2.1: Summary descriptive statistics

	Partial opening		**Entire opening**	
	mean	sd	mean	sd
Price (log)	11.08	0.86	10.86	0.87
MEA dummy	0.82	0.39	0.80	0.40
McEA dummy	0.11	0.31	0.12	0.32
MEA dummy (with the stations)	0.06	0.24	0.39	0.49
McEA dummy (with the stations)	0.01	0.10	0.03	0.16
Acreage (m^2)	0.05	0.16	0.05	0.29
Distance from station (m)	28.82	46.58	28.04	46.83
Building-area ratio	44.65	28.62	62.92	11.49
Floor-area ratio	225.58	118.19	219.59	113.87
Supply system dummies				
Water	0.99	0.12	0.99	0.10
Gas	0.49	0.50	0.50	0.50
Drain	0.66	0.47	0.74	0.44
Land use regulation dummies				
No regulation	0.13	0.34	0.13	0.34
Residential purpose	0.60	0.49	0.59	0.49
Commercial purpose	0.21	0.40	0.22	0.41
Industrial purpose	0.06	0.24	0.06	0.24
N	5488		4740	

Table 2.2: Characteristics of MEAs in Kyushu

MEA	Partial Opening	Entire Opening	Population (thousands)	Total area (ha)	Inhabitable land (ha)	Number of establishments	Number of employees
Kitakyushu MEA			1370.2	122139	67126	63842	654953
Fukuoka MEA		⋆	2495.6	128261	76431	109819	1260171
Omuta MEA		⋆	246.8	33222	25044	11120	97995
Kurume MEA		⋆	432.4	46841	35915	20422	192200
Iizuka MEA			196.5	38356	19273	8895	81356
Saga MEA			405.1	94134	59537	19578	191323
Nagasaki MEA			803.8	118499	58599	36459	364580
Sasebo MEA			304.6	55199	25237	14740	134210
Kumamoto MEA		⋆	1102.4	160364	97206	47198	505462
Yatsushiro MEA	⋆	⋆	145	71389	20918	7293	59197
Oita MEA			743.3	191335	65197	34027	352656
Miyazaki MEA			506.3	156073	57419	24572	234910
Miyakonojo MEA			243.6	123909	52532	12010	106413
Nobeoka MEA			131.2	86800	13669	6933	59020
Kagoshima MEA	⋆	⋆	731.5	103144	43381	35205	348729

The data source is Statistical Observations of Municipalities and Statistical Observations of Prefectures (The System of Social and Demographic Statistics of Japan) available in e-Stat. Population, Total area, and Inhabitable land show the values as of 2010. Number of establishments and employees show the values as of 2009. A ⋆ represents that a MEA/McEA had at least one Shinkansen station.

Table 2.3: Characteristics of McEAs in Kyushu

McEA	Partial Opening	Entire Opening	Population (thousands)	Total area (ha)	Inhabitable land (ha)	Number of establishments	Number of employees
Tagawa McEA			134.5	36365	14549	5920	46250
Yanagawa McEA			71.4	7688	7675	3364	25689
Karatsu McEA			133.3	52349	25624	6678	56560
Tosu McEA		⋆	113.1	14574	10414	4727	58104
Imari McEA			78.1	32082	14747	4423	37911
Shimabara McEA			97.8	25269	15292	5878	38523
Goto McEA			40.6	42085	14687	2741	15636
Hitoyoshi McEA			93.5	128362	26457	5012	38626
Minamata McEA	⋆	⋆	32	19688	5321	1678	14062
Tamana McEA		⋆	69.5	15255	12559	2863	24236
Yamaga McEA			66.6	39842	19169	3012	25679
Amakusa McEA			97.4	75034	24338	6203	39674
Nakatsu McEA			208.9	131619	44702	10642	92978
Hita McEA			70.9	66619	11244	4724	33007
Saiki McEA			77	90352	11668	4537	32979
Nichinan McEA			57.7	53612	11632	3227	24076
Hyuga McEA			88.3	90549	12792	4989	38200
Kanoya McEA			152.2	104814	40024	7644	62926
Makurazaki McEA			23.6	7488	4309	1393	10570
Satsumasendai McEA	⋆	⋆	154.8	109897	37266	7765	69155
Kirishima McEA			139.1	74801	23545	5802	62991
Amami McEA			54	47838	8560	3451	22928

The data source is Statistical Observations of Municipalities and Statistical Observations of Prefectures (The System of Social and Demographic Statistics of Japan) available in e-Stat. Population, Total area, and Inhabitable land show the values as of 2010. Number of establishments and employees show the values as of 2009. A ⋆ represents that a MEA/McEA had at least one Shinkansen station.

Table 2.4: Analysis 1

	Partial Opening			Entire Opening			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MEA (with the stations) × After	0.113***		0.113***	0.081***		0.081***	0.078***
	(0.028)		(0.028)	(0.023)		(0.023)	(0.023)
McEA (with the stations) × After		0.030	0.037		-0.038	-0.006	-0.004
		(0.027)	(0.027)		(0.026)	(0.024)	(0.024)
Covariates	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓
N	5487	5487	5487	4740	4740	4740	4678
R^2	0.811	0.811	0.811	0.829	0.828	0.829	0.831
adj. R^2	0.804	0.804	0.804	0.822	0.821	0.822	0.824

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(3) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (4)-(7) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (7) uses 2009 data instead of 2008 data in order to mitigate the effects of the Great Recession on land prices. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: Analysis 2

	Partial Opening			Entire Opening			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fukuoka MEA × After				0.107***		0.107***	0.107***
				(0.024)		(0.024)	(0.023)
Kurume MEA × After				0.057**		0.057**	0.058***
				(0.022)		(0.023)	(0.022)
Omuta MEA × After				-0.057*		-0.057*	-0.034
				(0.034)		(0.034)	(0.026)
Kumamoto MEA × After				0.085***		0.085***	0.079***
				(0.022)		(0.023)	(0.019)
Yatsushiro MEA × After	-0.079***		-0.079***	-0.067***		-0.067***	-0.092***
	(0.017)		(0.017)	(0.024)		(0.024)	(0.027)
Kagoshima MEA × After	0.131***		0.132***	0.071***		0.070***	0.051**
	(0.015)		(0.015)	(0.026)		(0.027)	(0.022)
Tosu McEA × After					-0.015	0.017	0.018
					(0.033)	(0.031)	(0.032)
Tamana McEA × After					-0.031	0.000	0.008
					(0.022)	(0.015)	(0.014)
Minamata McEA × After		-0.060***	-0.054***		-0.067***	-0.035***	-0.008
		(0.018)	(0.018)		(0.014)	(0.012)	(0.011)
Satsumasendai McEA × After		0.047*	0.054**		-0.063***	-0.031	-0.035*
		(0.026)	(0.025)		(0.023)	(0.022)	(0.020)
Covariates	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓
N	5487	5487	5487	4740	4740	4740	4678
R^2	0.811	0.811	0.811	0.829	0.828	0.829	0.831
adj. R^2	0.804	0.804	0.804	0.822	0.821	0.822	0.824

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(3) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (4)-(7) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (7) uses 2009 data instead of 2008 data in order to mitigate the effects of the Great Recession on land prices. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Analysis 3 (Partial opening in 2004)

NAME	MEA/McEA	0-5km	5-10km	10-15km	15-20km	20-25km	25km-
Yatsushiro	MEA	-0.085***	-0.130***	0.114*			
Kagoshima	MEA	0.293***	-0.010	-0.259***	0.339***	-0.167***	-0.830***
Minamata	McEA	-0.052***					
Satsumasendai	McEA	0.024	0.343***	-0.020	0.116***		

Standard errors are clustered in the municipalities. This shows the results for the opening in 2004. It uses data for 2001 and 2007. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.7: Analysis 3 (Entire opening in 2011)

NAME	MEA/McEA	0-5km	5-10km	10-15km	15-20km	20-25km	25km-
Fukuoka	MEA	0.312***	0.108***	-0.128*	0.033	-0.093	-0.086
Omuta	MEA	0.072**	-0.169	-0.183***			
Kurume	MEA	0.082***	-0.015	0.167***	0.535***		
Kumamoto	MEA	0.162***	0.115***	-0.098	-0.152**	0.212***	
Yatsushiro	MEA	-0.030	-0.339***	0.092			
Kagoshima	MEA	0.312***	-0.085***	-0.461***	0.095	-0.510***	-1.363***
Tosu	McEA	0.052***	-0.098***				
Minamata	McEA	-0.036***					
Tamana	McEA	0.098***	-0.188***				
Satsumasendai	McEA	0.094**	0.112**	-0.181*	-0.008		

Standard errors are clustered in the municipalities. This shows the results for the extension in 2011. It uses data for 2008 and 2014. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.8: Analysis3 (Entire opening in 2011, 2009-2014)

NAME	MEA/McEA	0-5km	5-10km	10-15km	15-20km	20-25km	25km-
Fukuoka	MEA	0.314***	0.099***	-0.134**	0.055	-0.041	-0.070
Omuta	MEA	0.104***	-0.137	-0.169***			
Kurume	MEA	0.083***	-0.015	0.160***	0.538***		
Kumamoto	MEA	0.156***	0.105***	-0.095	-0.104	0.239***	
Yatsushiro	MEA	-0.055***	-0.370***	0.059			
Kagoshima	MEA	0.290***	-0.111***	-0.480***	0.098	-0.511***	-1.385***
Tosu	McEA	0.055***	-0.098***				
Minamata	McEA	-0.008					
Tamana	McEA	0.108***	-0.187***				
Satsumasendai	McEA	0.088**	0.108**	-0.182*	-0.015		

Standard errors are clustered in the municipalities. This shows the results for the extension in 2011. It uses data for 2009 and 2014. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.9: Placebo test

	Partial opening			Entire opening		
	(1)	(2)	(3)	(4)	(5)	(6)
MEA (with the stations) × After	0.023		0.023	-0.034		-0.034
	(0.038)		(0.039)	(0.038)		(0.038)
McEA (with the stations) × After		-0.100*	-0.099		0.008	-0.006
		(0.060)	(0.061)		(0.025)	(0.028)
Covariates	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓
N	5089	5089	5089	5375	5375	5375
R^2	0.811	0.811	0.811	0.811	0.811	0.811
adj. R^2	0.804	0.804	0.804	0.804	0.804	0.804

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 1994 and 2000. Columns (5)-(8) are the results for the extension in 2011. They use data for 2002 and 2008. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: Sub-sample Analysis 1 by zoning

	Partial Opening				Entire Opening			
	(1) Full	(2) Residential	(3) Commercial	(4) Industrial	(5) Full	(6) Residential	(7) Commercial	(8) Industrial
MEA (with the stations) × After	0.113***	0.174***	0.044	-0.031	0.081***	0.068***	0.072***	0.020
	(0.028)	(0.035)	(0.063)	(0.035)	(0.023)	(0.023)	(0.025)	(0.039)
McEA (with the stations) × After	0.037	0.098*	0.034		-0.006	-0.003	0.038	0.005
	(0.027)	(0.050)	(0.054)		(0.024)	(0.027)	(0.050)	(0.042)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓	✓
N	5487	3313	1127	334	4740	2808	1020	277
R^2	0.811	0.788	0.840	0.781	0.829	0.787	0.872	0.764
adj. R^2	0.804	0.777	0.818	0.754	0.822	0.775	0.852	0.727

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (5)-(8) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (1) and (5) are full sample analyses. Columns (2) and (6) are sub-sample analyses of residential use. Columns (3) and (7) are sub-sample analyses of commercial use. Columns (4) and (8) are sub-sample analyses of industrial use. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: Sub-sample Analysis 2 by zoning

	Partial Opening				Entire Opening			
	(1) Full	(2) Residential	(3) Commercial	(4) Industrial	(5) Full	(6) Residential	(7) Commercial	(8) Industrial
Fukuoka MEA × After					0.107***	0.097***	0.070***	0.063*
					(0.025)	(0.024)	(0.025)	(0.035)
Kurume MEA × After					0.057**	0.028*	0.019	-0.045
					(0.023)	(0.015)	(0.028)	(0.053)
Omuta MEA × After					-0.057	-0.079***	0.026	-0.068**
					(0.035)	(0.024)	(0.079)	(0.031)
Kumamoto MEA × After					0.085***	0.080***	0.074***	0.060
					(0.023)	(0.016)	(0.027)	(0.038)
Yatsushiro MEA × After	-0.079***	-0.033	-0.133***		-0.067***	-0.136***	-0.011	
	(0.017)	(0.022)	(0.049)		(0.025)	(0.029)	(0.032)	
Kagoshima MEA × After	0.132***	0.196***	0.073	-0.031	0.070**	0.034	0.145***	-0.118***
	(0.015)	(0.022)	(0.052)	(0.036)	(0.027)	(0.029)	(0.041)	(0.032)
Tosu McEA × After					0.017	0.010	0.183***	0.005
					(0.031)	(0.049)	(0.021)	(0.046)
Tamana McEA × After						-0.012	0.038	
					(0.015)	(0.013)	(0.035)	
Minamata McEA × After	-0.054***	0.011	-0.007		-0.035***	0.004	-0.070***	
	(0.018)	(0.022)	(0.045)		(0.012)	(0.012)	(0.021)	
Satsumasendai McEA × After	0.054**	0.117*	0.046		-0.031	-0.018	-0.010	
	(0.026)	(0.059)	(0.065)		(0.022)	(0.019)	(0.049)	
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓	✓
N	5488	3315	1129	336	4740	2808	1020	281
R^2	0.811	0.788	0.841	0.782	0.829	0.788	0.873	0.771
adj. R^2	0.804	0.778	0.818	0.753	0.822	0.775	0.851	0.727

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (5)-(8) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (1) and (5) are full sample analyses. Columns (2) and (6) are sub-sample analyses of residential use. Columns (3) and (7) are sub-sample analyses of commercial use. Columns (4) and (8) are sub-sample analyses of industrial use. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Trains types on the Kagoshima route

		Train Types		
		Mizuho	Sakura	Tsubame
Hakata St.	Fukuoka MEA	●	●	●
Shin-Tosu St.	Tosu McEA		●	●
Kurume St.	Kurume MEA		●	●
Chikugo-Funagoya St.	Kurume MEA			●
Shin-Omuta St.	Omuta MEA			●
Shin-Tamana St.	Tamana McEA			●
Kumamoto St.	Kumamoto MEA	●	●	●
Shin-Yatsuhshiro St.	Yatsushiro MEA		○	●
Shin-Minamata St.	Minamata McEA		○	●
Izumi St.			○	●
Sendai St.	Satsumasendai McEA		●	●
Kagoshima-Chuo St.	Kagoshima MEA	●	●	●

●: served by all trains, ●: served by some trains. Source: Kyushu Railway Company Timetable (access: 2018/8/7) (https://www.jrkyushu-timetable.jp/cgi-bin/sp/sp-tt_dep.cgi/2862600/)

Table 2.12: Dynamic effects

	Partial opening		Entire opening	
	MEA dummy with the stations	McEA dummy with the stations	MEA dummy with the stations	McEA dummy with the stations
(base year + 1 year) dummy	0.023***	-0.013	0.028*	-0.004
	(0.003)	(0.020)	(0.015)	(0.005)
(base year + 2 year) dummy	0.039***	0.011	0.031**	-0.006
	(0.007)	(0.019)	(0.014)	(0.008)
(base year + 3 year) dummy	0.043***	0.026	0.028**	-0.007
	(0.013)	(0.018)	(0.014)	(0.012)
(base year + 4 year) dummy	0.078***	0.065***	0.035*	0.004
	(0.017)	(0.023)	(0.019)	(0.023)
(base year + 5 year) dummy	0.084***	0.018	0.049**	-0.003
	(0.022)	(0.033)	(0.022)	(0.022)
(base year + 6 year) dummy	0.112***	0.033	0.080***	-0.005
	(0.026)	(0.035)	(0.029)	(0.023)
(base year + 7 year) dummy	0.137***	0.023	0.109***	-0.010
	(0.028)	(0.037)	(0.036)	(0.027)
(base year + 8 year) dummy	0.142***	0.019	0.131***	-0.016
	(0.031)	(0.037)	(0.042)	(0.031)
(base year + 9 year) dummy	0.163***	0.018	0.150***	-0.023
	(0.033)	(0.039)	(0.051)	(0.037)
N	27355	27355	23742	23742
R^2	0.810	0.810	0.828	0.828
adj. R^2	0.809	0.809	0.827	0.827

Standard errors in parentheses are clustered in the municipalities. The analysis for the opening in 2004 uses data from 2000 to 2009. The analysis for the extension in 2011 uses data from 2007 to 2016. The years 2000 and 2007 are the base years for the partial opening and the whole opening, respectively. Each value shows the estimated coefficient of MEA (or McEA) with the Shinkansen stations dummy times each year dummy. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.14: Analysis 1 with train type dummies

	Entire opening	
	(1)	(2)
MEA dummy (with the stations) × After	0.081***	-0.059**
	(0.023)	(0.029)
McEA dummy (with the stations) × After	-0.006	-0.015
	(0.024)	(0.018)
three types dummy × MEA dummy (with the stations) × After		0.154***
		(0.030)
two types dummy × MEA dummy (with the stations) × After		0.116***
		(0.032)
two types dummy × McEA dummy (with the stations) × After		0.010
		(0.028)
Covariates	✓	✓
Time F.E.	✓	✓
Municipalities F.E.	✓	✓
N	4740	4740
R^2	0.829	0.829
adj. R^2	0.822	0.822

Standard errors in parentheses are clustered in the municipalities. All columns show the results for the extension in 2011. They use data for 2008 and 2014. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.8 Appendices

Appendix A: Results for the Control Variables

	Partial Opening			Entire Opening		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	9.592***	9.602***	10.371***	9.853***	9.900***	10.412***
	(0.190)	(0.197)	(0.135)	(0.167)	(0.176)	(0.160)
Distance from the nearest station	-0.003***	-0.003***	-0.004***	-0.003***	-0.003***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Acreage	-0.293***	-0.291***	-0.112	-0.112***	-0.109***	-0.090
	(0.077)	(0.078)	(0.091)	(0.022)	(0.022)	(0.098)
Building-area ratio	-0.002***	-0.002***	-0.005***	-0.010***	-0.010***	-0.011***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Floor-area area	0.002***	0.002***	0.003***	0.003***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Water	0.264**	0.265*	-0.497***	0.212*	0.207	-0.002
	(0.133)	(0.138)	(0.045)	(0.127)	(0.134)	(0.122)
Gas	0.305***	0.309***	0.401***	0.304***	0.308***	0.264***
	(0.033)	(0.033)	(0.099)	(0.030)	(0.031)	(0.043)
Drain	0.195***	0.190***	0.299***	0.176***	0.176***	0.188**
	(0.027)	(0.028)	(0.051)	(0.034)	(0.036)	(0.073)
Residential purpose	0.772***	0.785***	0.762***	0.632***	0.614***	0.592***
	(0.065)	(0.069)	(0.128)	(0.043)	(0.043)	(0.038)
Commercial purpose	1.221***	1.202***	0.988***	1.000***	0.943***	0.813***
	(0.049)	(0.053)	(0.143)	(0.064)	(0.064)	(0.081)
Industrial purpose	0.813***	0.820***	0.590***	0.553***	0.528***	0.449***
	(0.066)	(0.071)	(0.138)	(0.052)	(0.052)	(0.074)
N	5487	5095	396	4740	4370	1846
R^2	0.811	0.803	0.856	0.829	0.823	0.853
adj. R^2	0.804	0.797	0.844	0.822	0.817	0.845

Standard errors in parentheses are clustered in the municipalities. Column (1) - (3) are the results for the opening in 2004. They use data for 2001 and 2007. Column (4) - (6) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (1) and (4) show the estimated covariates of Columns (3) and (6) in Table 2.4, respectively. Column (2) and (5) show the estimated covariates of Column (3) and (6) in Table 2.5, respectively. Column (3) shows the result of covariates in Table 2.6. Column (6) shows the result of covariates in Table 2.7. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix B: Relationship between land price and distance from the Shinkansen station within a city

Figure 2.5: Relationship between land price and distance from the Shinkansen station in 4 MEAs

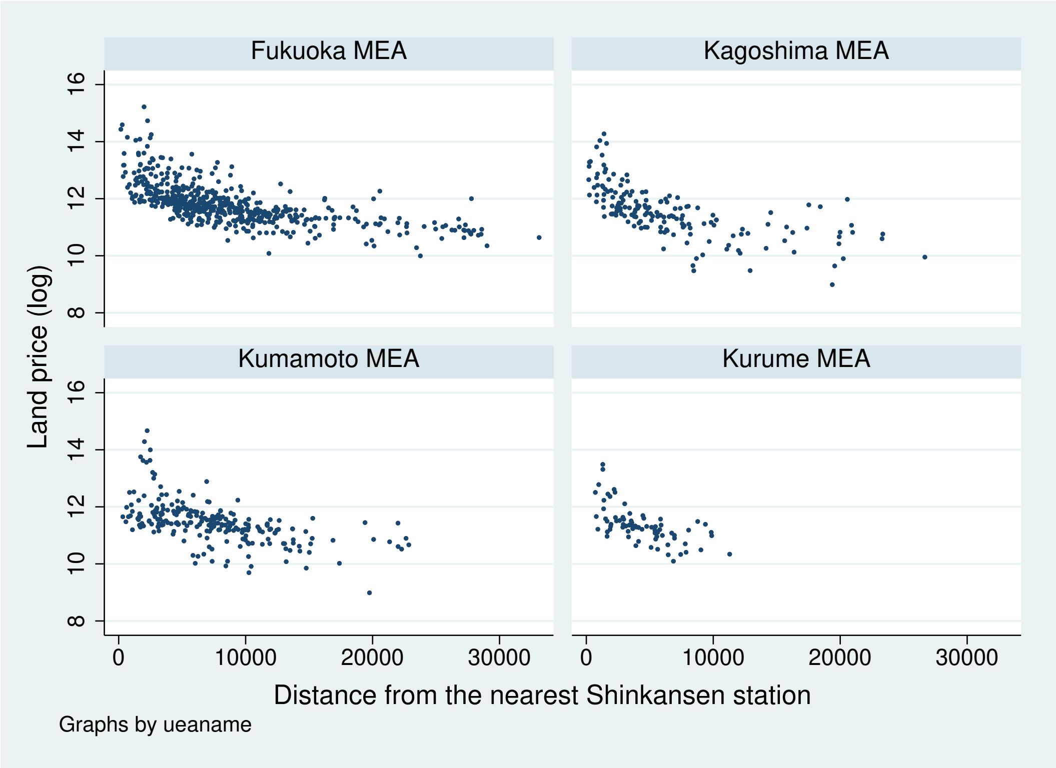
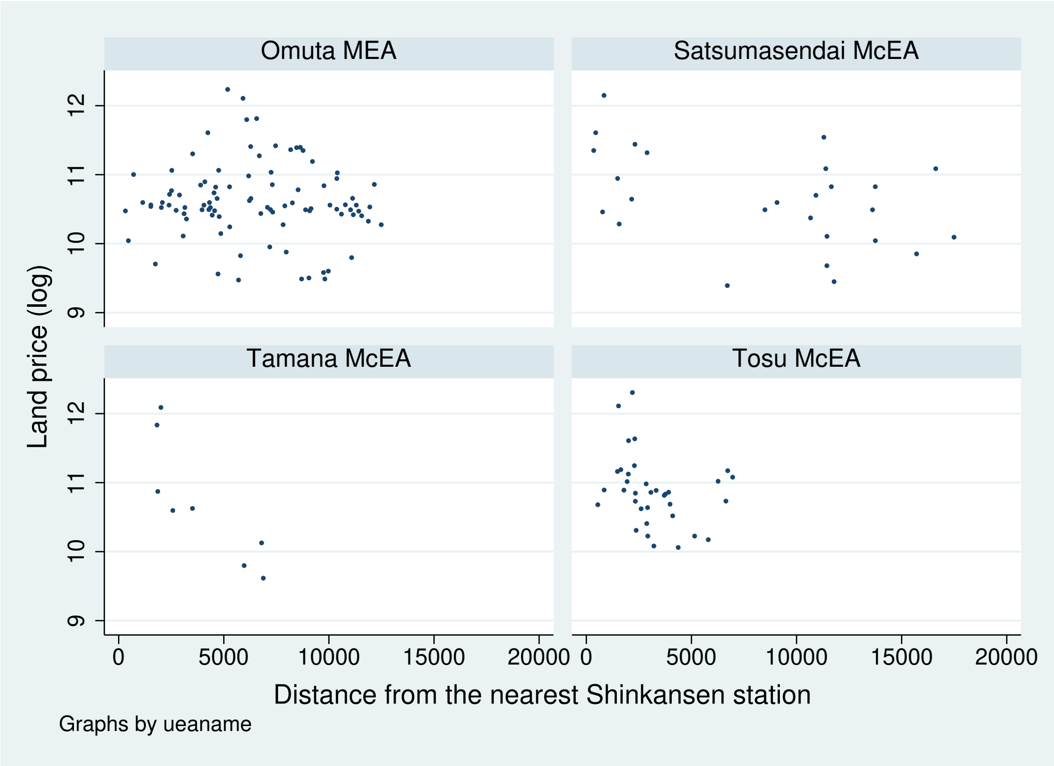


Figure 2.6: Relationship between land price and distance from the Shinkansen station in 1 MEA and 3 McEAs



Appendix C: Robustness check against exclusion of Saga, Nagasaki, and/or Kitakyusyu

Table 2.15 and Table 2.16 show the robustness check for analysis 1 and analysis 2, respectively. For each table, columns (1)-(4) show the results for the partial opening in 2004 and columns (5)-(8) show the results for the opening of the entire route in 2011. Columns (1) and (5) in Table 2.15 (resp. Table 2.16) are the baseline results that come from columns (3) and (6) in Table 2.4 (resp. Table 2.5). Columns (2) and (6) excluded points in Kitakyushu MEA , and columns (3) and (7) do points in Saga prefecture and Nagasaki prefecture. Finally, columns (4) and (8) exclude all the points in Saga prefecture, Nagasaki prefecture and Kitakyushu MEA. Those results show that our main results are unaltered.

Table 2.15: Robustness check for analysis 1

	Partial Opening				Entire Opening			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MEA dummy (with the stations) × After	0.113***	0.109***	0.119***	0.114***	0.081***	0.092***	0.077***	0.092***
	(0.028)	(0.029)	(0.030)	(0.032)	(0.023)	(0.023)	(0.025)	(0.026)
McEA dummy (with the stations) × After	0.037	0.032	0.037	0.032	-0.006	0.005	-0.032	-0.017
	(0.027)	(0.028)	(0.027)	(0.029)	(0.024)	(0.024)	(0.020)	(0.021)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓	✓
N	5487	4800	4546	3859	4740	4148	3926	3334
R^2	0.811	0.814	0.833	0.840	0.829	0.833	0.842	0.849
adj. R^2	0.804	0.807	0.827	0.834	0.822	0.825	0.836	0.842

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (5)-(8) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (1) and (5) include full sample. Columns (2) and (6) exclude points in Kitakyushu MEA. Columns (3) and (7) exclude points in Saga prefecture and Nagasaki prefecture. Columns (4) and (8) exclude all the points in Saga prefecture, Nagasaki prefecture and Kitakyushu MEA. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.16: Robustness check for analysis 2

	Opening				Extension			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fukuoka MEA × After					0.107***	0.118***	0.104***	0.120***
					(0.025)	(0.025)	(0.026)	(0.028)
Kurume MEA × After					0.057**	0.069***	0.049*	0.065**
					(0.023)	(0.023)	(0.026)	(0.027)
Omuta MEA × After					-0.057	-0.047	-0.061*	-0.048
					(0.035)	(0.035)	(0.037)	(0.037)
Kumamoto MEA × After					0.085***	0.097***	0.081***	0.097***
					(0.023)	(0.023)	(0.025)	(0.026)
Yatsushiro MEA × After	-0.079***	-0.084***	-0.081***	-0.088***	-0.067***	-0.063**	-0.057*	-0.053
	(0.017)	(0.020)	(0.017)	(0.019)	(0.025)	(0.027)	(0.033)	(0.034)
Kagoshima MEA × After	0.132***	0.128***	0.139***	0.133***	0.070**	0.080***	0.065**	0.078**
	(0.015)	(0.017)	(0.017)	(0.020)	(0.027)	(0.028)	(0.028)	(0.030)
Tosu McEA × After					0.017	0.028		
					(0.031)	(0.031)		
Tamana McEA × After					0.000	0.012	-0.007	0.009
					(0.015)	(0.016)	(0.019)	(0.020)
Minamata McEA × After	-0.054***	-0.059***	-0.057***	-0.064***	-0.035***	-0.024*	-0.038**	-0.022
	(0.018)	(0.020)	(0.018)	(0.020)	(0.012)	(0.013)	(0.015)	(0.017)
Satsumasendai McEA × After	0.054**	0.050*	0.056**	0.051*	-0.031	-0.021	-0.038	-0.023
	(0.026)	(0.027)	(0.026)	(0.027)	(0.022)	(0.023)	(0.023)	(0.025)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓	✓
N	5488	4801	4547	3860	4740	4148	3926	3334
R^2	0.811	0.814	0.833	0.840	0.829	0.833	0.843	0.849
adj. R^2	0.804	0.807	0.827	0.834	0.822	0.825	0.836	0.842

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (5)-(8) are the results for the extension in 2011. They use data for 2008 and 2014. Columns (1) and (5) include full sample. Columns (2) and (6) exclude points in Kitakyushu MEA. Columns (3) and (7) exclude points in Saga prefecture and Nagasaki prefecture. Columns (4) and (8) exclude all the points in Saga prefecture, Nagasaki prefecture and Kitakyushu MEA. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix D: Detail sub-sample analysis by land-use zoning

In the sub-sample analysis conducted in section 6.2., we divide twelve categories of land use zoning into three groups: residential, commercial, and industrial uses. Among residential land use, there are seven categories and some categories can have stores or offices depending on their floor space. In this appendix, we conduct more detailed sub-sample analysis especially for residential land use in order to examine which types of agents (residents, stores, or offices) have been more affected by the high-speed railway. Seven categories of residential land use are following: (i) Category I Exclusively Low-rise Residential Zone, (ii) Category II Exclusively Low-rise Residential Zone, (iii) Category I Mid/high-rise Oriented Residential Zone, (iv) Category II Mid/high-rise Oriented Residential Zone, (v) Category I Residential Zone, (vi) Category II Residential Zone, and (vii) Quasi-residential zone. Here, we use specification model (1) (i.e., analysis 1). Columns (1)-(4) show the results for the partial opening in 2004 and columns (5)-(8) show the results for the opening of the entire route in 2011. Columns (1), (2), (3), and (4) (and columns (5), (6), (7), and (8)) show the results of sub-sample analyses for category (i), categories (i) and (ii), categories (iii) and (iv), and categories (v) - (vii).

According to the land-use regulation, category (i) can not have any stores. Categories (ii) - (vii) can have stores with 150m^2 or less, 500m^2 or less, 1,500m^2 or less, 3,000m^2 or less, 10,000m^2 or less and 10,000 m^2 or less with some other conditions. In contrast, categories (i) - (iii) can not have any offices. Category (iv) can have offices with 1,500 m^2 or less and 2 floors or less. Category (v) can have offices with 30,000 m^2 or less and categories (vi) - (ii) can have offices over 30,000m^2.

Columns (1), (2), (5), and (6) in Table 2.17 show that points in categories (i) or (ii) of treated MEAs have significantly positive effects of partial opening but they have insignificant or less significant effects of entire opening. It implies that partial opening affect residents more but entire opening affect residents less. Column (3), (4), (7), and (8) show that both partial opening and entire opening have positive impacts on points in categories (iii) - (vii) of treated MEAs or McEAs.

Table 2.17: Detail sub-sample analysis by zoning

	Partial Opening				Entire Opening			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MEA (with the stations) × After	0.213***	0.206***	0.092	0.135***	0.038**	0.022	0.096***	0.080**
	(0.027)	(0.026)	(0.076)	(0.030)	(0.019)	(0.017)	(0.026)	(0.039)
McEA (with the stations) × After			0.171***	-0.006	-0.028	-0.042	-0.031	0.015
			(0.048)	(0.041)	(0.061)	(0.055)	(0.032)	(0.026)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓	✓	✓
Municipalities F.E.	✓	✓	✓	✓	✓	✓	✓	✓
N	860	985	886	1431	705	804	735	1257
R^2	0.782	0.793	0.830	0.824	0.764	0.778	0.838	0.829
adj. R^2	0.762	0.773	0.810	0.806	0.737	0.751	0.816	0.808

Standard errors in parentheses are clustered in the municipalities. Columns (1)-(4) are the results for the opening in 2004. They use data for 2001 and 2007. Columns (5)-(8) are the results for the extension in 2011. They use data for 2008 and 2014. Seven categories of residential land use are following: (i) Category I Exclusively Low-rise Residential Zone, (ii) Category II Exclusively Low-rise Residential Zone, (iii) Category I Mid/high-rise Oriented Residential Zone, (iv) Category II Mid/high-rise Oriented Residential Zone, (v) Category I Residential Zone, (vi) Category II Residential Zone, and (vii) Quasi-residential zone. Columns (1), (2), (3), and (4) (and columns (5), (6), (7), and (8)) show the results of sub-sample analyses for Category (i), categories (i) and (ii), categories (iii) and (iv), and categories (v) - (vii). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

Neighborhood externalities from "one-room apartments"

3.1 Introduction

Policymakers have used externalities to justify government intervention (Gerardi et al., 2015). A recent example in Japan is the regulations around "one-room apartments" implemented by several urban municipalities such as Tokyo's 23 wards. One-room apartments here refer to studio apartments that consist of many "one-room units," which are small housing units mainly for single-person households. These apartments are proliferating, especially in Tokyo, because of the dramatic increase in single-person households. Alongside this increase, the negative externalities of residents in one-room apartments, especially one-room units, are rising. For example, neighborhood residents often complain that one-room residents create issues in terms of noise, waste disposal, and street parking. It is also said that these apartments are maintained to a lower standard because their owners have bought them for investment purposes rather than to live in themselves. In response, several municipalities have regulated the construction of new one-room apartments, although they have not clearly mentioned the existence of the above mentioned externalities.[1] This study investigates whether such policies are effective as well as whether one-room residents cause such externalities by using a rich dataset of rentals in Tokyo's 23 wards after 2000. By estimating the hedonic price equations of non-one-room units (family units hereafter), we examine whether externalities of one-room units toward family units within an apartment or district exist.

The dataset used in this study contains detailed information on each rental including rent per month, attributes, registered date, and apartment block. With this information, we specify for each apartment (1) whether the apartment has one-room units within, (2) which floors have one-room units and (3) how many one-room units the apartment has. Then, we test the existence of negative externalities of one-room units by regressing the indicators of one-room units on the rents of family units. Apartment or district fixed effects are used in the regressions and several sub-analyses are conducted to reduce the bias caused by omitted variables such as the unobservable quality differences among apartments. We also examine whether negative externalities of one-room units within a district exist by using the share of households living in one-room units in each district as the main regressor.

This study is related to the large literature on the impact of neighborhood externalities on housing rents or prices. Much scholarly attention has been paid to environmental externalities such as air quality and water quality (Boyle et al., 2001). Recent studies have also found that train or air noise has negative effects on property values (Diao et al., 2016; Winke, 2016). Another strand of the literature has focused on the externalities of specific housings. For example, several studies have addressed the negative price externalities of foreclosure (Gerardi et al.,

[1] Many municipalities of Tokyo's 23 wards are planning regulations to build better neighborhood relationships and prevent the conflicts caused by the construction of one-room apartments.

2015; Zhang and Leonard, 2014). Revitalization programs of housings have also been found to positively affect the neighborhood (Leonard et al., 2017; Rossi-Hansberg et al., 2010; Sandler, 2017). The recent study by Shimizu and Nakagawa (2018) in Japan presented the neighborhood negative externalities of deteriorated apartments. The other strand of the literature has focused on the social status of neighbors. For example, Ioannides and Zabel (2008) showed that homeowners prefer to live with neighbors with similar characteristics and Leung and Tsang (2012) found that homeowners have a strong distaste for income inequality in their neighborhood. All these studies analyze the neighborhood externalities within an area, district, or region. On the other hand, this study focuses on the externalities within an apartment besides the externalities within a district. Thus, we contribute to the literature by studying the most micro-level neighborhood effect, which has not thus far been investigated to the best of our knowledge. Liu et al. (2018) mentioned that the vertical pattern of urban development has been largely ignored and estimated the vertical rent gradient. Although they focused on commercial real estate, the same argument can be applied to residential real estate. While previous studies of neighborhood effects have focused only on "horizontal externalities" (i.e., externalities toward horizontal direction), we also examine "vertical externalities" (i.e., externalities toward vertical direction) within an apartment.

The main results are as follows. Firstly, we focus on the apartments which have both family and one-room units (mixed apartments, hereafter). Family units that have one-room units on the same floor have lower rents than family units that do not have them, ceteris paribus. More specifically, a 1 percentage point increase in the share of one-room units within a mixed apartment decreases the rents of family units by around 0.05%. These results support the evidence that negative externalities of one-room units within an apartment exist. We also find that smaller one-room units generate externalities more. Secondly, we focus on the family units in any type of apartments. By using the district-level share of one-room residents as the main regressor, we find that increases in this share lower the rents of family units, implying that externalities within a district also exist. In conclusion, this study provides evidence of the vertical and horizontal neighborhood externalities of one-room units.

The remainder of this chapter is structured as follows. Section 3.2 describes one-room apartments in Japan and the neighborhood externalities of these apartments. Section 3.3 explains the data and Section 3.4 presents our specification model. Section 3.5 shows the estimation results and Section 3.6 concludes.

3.2 Research background

One-room units are housing units for single-person households that have a main room, a bathroom, and a kitchen.[2] These apartments were built from the early 1980s to accommodate the large influx of single-person households into metropolitan areas (Ezawa and Nakagawa, 2009). As Figure 3.1 shows, the number and proportion of single-person households have been growing in Tokyo's 23 wards. In 2015, the proportion of single-person households in Tokyo's 23 wards reached approximately 50%. This increase in single-person households has led to a rise in the supply of one-room apartments. According to the Housing and Land Survey in Japan, the

[2]In this study, "housing units" or "units" refer to living spaces for households and "apartments" to apartments composed of several housing units.

number of housing units smaller than 30m^2 in Tokyo's 23 wards was 973,400 in 1998, but this number gradually increased to 1,077,600 in 2013.

[Figure 3.1]

As the number of one-room units has increased, the conflicts associated with the construction and management of one-room apartments have also grown, especially between residents/developers and neighborhoods (Kinoshita et al., 2008). According to Tonozuka (2010), neighborhood residents have two main complaints about the residents of one-room apartments: (1) bad manners in terms of noise, waste disposal, and street parking and (2) a lack of involvement in community activities. According to a questionnaire conducted by *Asahi Shimbun* (2008),[3] respondents believe that increases in the number of one-room apartments worsen the balance of household composition in the region. Further, some one-room apartments are constructed for investment and speculation in response to increasing demand for such small-scale residences. Hence, the owners of these apartments rarely live in the units themselves, which are therefore insufficiently managed.

In response to these conflicts, municipalities in Tokyo's 23 wards and other major cities including Osaka city, Nagoya city, Hakata city, and Kawasaki city have regulated one-room apartments to govern their construction and reconstruction. Although the content of the regulations varies by municipality, there are three major components: (1) regulation of minimum exclusive areas, (2) obligation to construct facilities such as parking areas, garbage dumps, and apartment management offices, and (3) obligation to include a certain proportion of housing units for families. These policies began to come into effect around 1984, soon after the beginning of the proliferation of one-room apartments. The initial policies were simply guidelines; however, in the 2000s, many municipalities began to implement legally binding and more restrictive policies. In 2018, 19 out of Tokyo's 23 wards had such ordinances. However, because these policies are similar regardless of whether they are guidelines or ordinances, we assume that their effect is common across Tokyo's 23 wards.[4]

3.3 Data

We use the Real Estate Database 1999–2016 provided by At Home Co., Ltd., which is one of the major real estate information providers in Japan. The data are based on information registered for advertising by real estate stores that are members of the information network of At Home. In particular, we use rental property data on Tokyo's 23 wards. Our observations are rental units and each unit has information including unit ID, apartment ID, rent per month, attributes of the unit and apartment, registered date, and whether the unit is second hand or not.

In this study, we classify units into two groups based on the size of floor space: *one-room units* and *family units*. In general, the former are rooms for single-person households and the latter are rooms for multiple-person households. Several acts and ordinances in Japan categorize units by floor space. "The Outline of the Basic Act for Housing" sets the minimum and target

[3] *Asahi Shimbun* is a major newspaper in Japan. This questionnaire was conducted on September 2018 in the membership service of *Asahi Shimbun*. There were about 16,200 responses.

[4] Kinoshita et al. (2008), Ezawa and Nakagawa (2009), and Tonozuka (2010) all reviewed the details of such one-room apartment regulations.

levels of floor space depending on the number of people per household.[5] Table 3.1 shows the shares of households that satisfied these levels in Tokyo's 23 wards as of 2013. In total, 61.1% of single-person households live below the target level (40m^2) and 26.7% of those households live below the minimum level (25m^2). Although the level differs slightly among municipalities, the majority of municipalities set one-room units as units smaller than 30m^2 or 40m^2 and family units as units larger than 40m^2 or 50m^2. This study thus defines units smaller than 40m^2 as one-room units and units equal to or greater than 40m^2 as family units. The analyses presented later only use family units as observations to analyze the neighborhood externalities of one-room units toward family units.[6]

[Table 3.1]

We also categorize apartments into three types: *one-room*, *mixed*, and *family* apartments. Family apartments (resp. one-room apartments) consist of only family units (resp. one-room units), while mixed apartments consist of both one-room units and family units. Table 3.2 shows the frequency and share of each apartment type in Tokyo's 23 wards. The share of mixed apartments is the largest among the three types. The analyses of the externalities within an apartment use family units in mixed apartments, while the analyses of the externalities within a district use family units in both mixed and family apartments.

[Table 3.2]

There are three main regressors. The first is the mixed dummy that takes one if the family unit has any one-room units on the same floor. The second one is the share of one-room units within the apartment in which the unit is located. Since the original data do not contain a variable on the number of units or one-room units in each apartment, we construct them by using unit ID and apartment ID. The third one is the share of one-room units in each district from the Population Censuses in 2000 and 2010. District is the finest unit in the dataset we use.[7] According to the 2015 Population Census, there are 3,142 districts in Tokyo's 23 wards. Their average area is 0.20 km^2, average population is 2,951, and average number of households is 1,528. The data details and data construction methods are provided in Appendix A.

As to the control variables, the analyses uses the unit's and apartment's attributes. The unit's attributes are the floor of the unit and floor space. The apartment's attributes are the number of floor above the ground, walking time to station, age, structural material, and nearest railway line.[8] Appendix B summarizes the summary statistics of the samples.

3.4 Specification model

Here, we examine the existence of negative externalities of one-room units within an apartment and then focus on the negative externalities within a district. In the analyses of apartments, we especially focus on mixed apartments to eliminate quality difference between mixed and family

[5] These levels are not legally binding.

[6] Generally, families living in apartments have lower income level than families living in privately owned houses. Since we limit the data sample to apartments, our analyses may underestimate the effect of the externalities.

[7] District refers to "Cho-choji tou" in Japanese.

[8] We classify the railway lines which share is less than 2% in the original data as "others."

apartments. In all the estimations, the unit of observation is a family unit. The dependent variable is $\ln(p_{it})$, which denotes the logarithm of the rent (in yen) of family unit i on floor l within apartment j located in area k at time t. We denote unit i's attribute vector by X_i, apartment j's attribute vector by Z_j, the year-month dummy by q_t, and the idiosyncratic error term by u_{it}. β, δ and ψ are the coefficients of the main independent variable related to one-room units, X_i and Z_j, respectively.

In the first analysis, we estimate whether the existence of one-room units decreases the rents of family units on the same floor. We call the floor on mixed apartments, which has both family and one-room units, as a "mixed floor." If negative externalities of one-room units toward family units exist, family units on a mixed floor would have lower rents than family units on a non-mixed floor, ceteris paribus. The specification model is as follows:

$$\ln(p_{it}) = \alpha + \beta \mathrm{mix}_l + X_i\delta + q_t + s_j + u_{it} \tag{3.1}$$

We are interested in mix_l, which is the dummy variable that takes one if floor l is a mixed floor. The identification of the estimator of interest requires no correlation between mix_l and the error term. The concern about this requirement is whether omitted variables regarding the quality of apartments, besides the observed attributes, exist. To control for such bias, we include the fixed effects of apartment s_j. It enables us to control for the time-invariant characteristics of each apartment, including unobserved qualities. It is conjectured that family units located next to or near the one-room units have a lower quality than other family units. To remove the rent difference caused by this quality variance, we use the attributes of the unit X_i. In sum, we estimate the effect of one-room units on family units using within-apartment variations. In other words, we compare family units on mixed floors with family units on non-mixed floors in the same apartment while controlling for each unit's attributes. If mix_l is significantly negative, family units on mixed floors have lower rents than family units on non-mixed floors, ceteris paribus, implying that negative externalities of one-room units exist. In the analyses, we eliminate the observations in the apartments which do not have multiple floors in the data.

In the second specification model, we estimate how many percent of the rents of family units get decreased if the share of one-room units in a mixed apartment gets increased by 1 percentage point. Same as the first specification model, the sample unit is a family unit in the mixed apartments. Here, we are interested in the share of one-room units in apartment j, share_j. The specification is as follows:

$$\ln(p_{it}) = \alpha + \beta \mathrm{share}_j + X_i\delta + Z_j\psi + q_t + s_k + u_{it} \tag{3.2}$$

Since the share of one-room units in each apartment has not changed over time, the second specification model should not involve the fixed effects of an apartment. Instead, the above specification model uses the fixed effects of a district, assuming that apartments within the same district may have similar characteristics in terms of apartment quality. We also include the attributes of apartments in order to control for the quality of each apartment.

The specification (3.2) has a fear of not sufficiently controlling for the quality difference among apartments. Since it is conjectured that apartments with more one-room units have lower quality, the estimated coefficient may include the effect of quality difference on rents. As a robustness check, we use the post-double-selection (PDS) methodology developed by Belloni et al. (2014). In order to avoid omitted variable problems, it is important to use sufficient and appropriate control variables. In the above specification, we use several variables as controls, which have a few missing values in original data, and exclude the observations with any missing value. However, there are other candidates of controls, such as interaction terms between variables or quadratic terms of each variable. If the values are missing, not randomly but systematically, excluding the observations with missing values generates bias in the estimation. The question of whether the value is missing or not is another candidate of controls. Although using all the candidates may ease the omitted variable problem, too many control variables cause overfitting problems. The PDS methodology employs the lasso estimator to choose optimal control variables from high-dimensional data. Lasso regression minimizes the mean squared error with ℓ_2 penalty, which shrinks all the coefficients toward zero. It forces some of the coefficients to be exactly equal to zero and thus, performs variable selection (James et al., 2013).

The estimation steps in the PDS methodology are as follows: (1) run the lasso regression with share_j as the dependent variable, and X_i and Z_j as the independent variables; (2) run the lasso regression with $\ln(p_{it})$ as the dependent variable, and X_i and Z_j as the independent variables; and (3) run the OLS regression with $\ln(p_{it})$ as the dependent variable, share_j as the main independent variable, and the union of selected variables in step 1 and step 2 as the other independent variables (i.e., the control variables).[9] According to Belloni et al. (2014), The first step aims to find control variables that are strongly related to share_j. These variables are potentially important confounding factors. The second step aims to find significant variables to predict $\ln(p_{it})$, which enables us to keep the residual variance small and find additional significant confounds. In this robustness check, new variables are added to controls. [10][11]

We now leave the neighborhood externalities of one-room units within an apartment and instead focus on the externalities within a district. Here, we examine whether the rents of family units increase as the share of households living in one-room units within a district rises over time:

$$\ln(p_{it}) = \alpha + \beta \text{share}_{kt} + X_i\delta + q_t + s_j + u_{it} \tag{3.3}$$

The unit of observation is a family unit in both mixed and family apartments. The dependent variables and the unit's attributes are the same as those in the first specification. However, we use the share of one-room units in each district in year t instead of that for each apartment. The data sources of the new main independent variable are the Population Censuses in 2000

[9] The district fixed effects and year-month fixed effects are included and unpenalized by lasso. As an exception, the year-month fixed effects in Table 3.6 are penalized because of the data limitation.

[10] Besides the unit's and apartment's attributes used in the basic analysis, controls in the PDS methodology include other variables with many missing values such as the total number of floor under the ground, balcony floor space, bus riding time, walking time to bus station, and a second-hand dummy. We create a missing value dummy for each variable. Quadratic terms of numeric variables and interaction terms between variables are also used as new controls.

[11] We use Stata command pdslasso (Ahrens et al., 2018).

and 2010. These provide the number of households living in a certain category of floor space in each district. The share of one-room units in district k is calculated by dividing the number of households living in a unit with a floor space of less than 30m^2 in district k by the number of households in district k. The rental data contain every year from 2000 to 2016, while we have only two years of data for the district-level share of one-room units. Thus, we only use units registered in 2001, 2002, 2003, 2004, 2005, 2011, 2012, 2013, 2014, and 2015 as our observations. We assign the district-level share of one-room units in 2000 (resp. 2010) to units in 2001 to 2005 (resp. 2011 to 2015). Here, we use the apartment fixed effects as in the first specification model, thus we can control unobservable quality differences of apartments.

3.5 Results

3.5.1 Externalities within an apartment

One-room units on the same floor

We focus on the existence of neighborhood externalities within a mixed apartment. Table 3.3 reports the estimation results based on specification (3.1). If the main regressor, the mixed floor dummy, is significantly negative, family units on mixed floors have lower rents than family units on non-mixed floors, ceteris paribus, indicating that negative neighborhood externalities of one-room units exist. The mixed apartment dummy in Panel A (resp. Panel B) of Table 3.3 equals one if at least one one-room unit smaller than 40m^2 (resp. 25m^2) exists on the same floor.

[Table 3.3]

The main regressor is significantly negative for all the columns in both panels. Column (1) in Panel A (resp. Panel B), which does not include the controls, suggests that family units on mixed floors have 4.4% (resp. 2.2%) lower rents than those on non-mixed floors. As columns (2)–(6) show in both panels, the value falls below 1% once we include the controls. In general, family units on mixed floors have lower structural quality than family units on non-mixed floors. Thus, we remove the effect of quality difference on rents by including the controls. We find that the absolute value of the estimated coefficient dramatically declines. Even if we control for the quality difference, the main regressor is significantly negative, implying that negative neighborhood externalities of one-room units within an apartment exist.

To remove the quality difference more robustly, column (3) excludes units in districts designated as "densely populated urban districts significantly at risk in the event of earthquakes and other natural disasters (the Ministry of Land, Infrastructure, Transport and Tourism)."[12] These districts are "urban districts with high potential for fire-spreading and/or extensive difficulty of evacuation, poising major risks or seriously undermining the securing of minimum levels of safety from among densely populated districts (the Ministry of Land, Infrastructure, Transport and Tourism)."[13] Column (4) only uses units built in 1982 or after, when the new standard for

[12] This designation was announced in October 2012 by the Ministry of Land, Infrastructure, Transport and Tourism, Fire and Disaster Management Agency, Ministry of Economy, Trade and Industry, and Cabinet Office.

[13] The data source is National Land Numerical Information Populated Urban Districts Data. Reference: http://nlftp.mlit.go.jp/ksj-e/gml/datalist/KsjTmplt-A39.html (accessed on February 12, 2019).

earthquake-resistant housing design began to be applied to newly built apartments.[14] The new standard demands stricter quality for newly built apartments than before. Columns (3) and (4) aim to omit the apartments which may have different building structures. To control the apartment and municipality specific time effects, columns (5) and (6) use the apartment-year and municipality-year fixed effects, respectively. According to Table 3.3, the estimated coefficients of columns (3) – (6) are similar to those of column (2).

Share of one-room units within an apartment

Next, we investigate the effect on the rents of family units if the share of one-room units within an apartment increases. The main regressor in Panel A (resp. Panel B) of Table 3.4 is the share of one-room units below 40m^2(resp. 25m^2). This analysis uses only family units in mixed apartments. Column (1) does not include the controls, whereas columns (2)–(6) do.

[Table 3.4]

The main regressor is significantly negative for all the columns. Column (1), which does not include the controls, suggests that a 1 percentage point increase in the share of one-room units raise rents by 0.092% (resp. 0.082%) in panel A (resp. panel B). As columns (2)–(6) show, the value falls once we include the controls. In general, mixed apartments have lower structural quality than family apartments. By including the controls, we can thus remove the effect of structural quality on rents and then the absolute value of the estimated coefficient dramatically declines. Even if we control for the quality difference, the main regressor is significantly negative, indicating that negative neighborhood externalities of one-room units within an apartment exist.

Column (3) excludes units in densely populated urban districts to eliminate those with lower quality. Column (4) includes the time trend in each municipality to control the effect of each municipality's policy such as the one-room regulation. More finely, column (5) includes the time trend in each area. Although the fixed effects of district control for the time-invariant effects in each district, some districts may have witnessed increasing demand during the sample period because of the introduction of new facilities or trends. Column (5) aims to control for this time-variant effect in each district. Finally, column (6) conducts the PDS estimation described in the previous subsection. Columns (3)–(6) have the same sign of each estimated coefficient as column (2), and the value of each coefficient also changes little among them.

In Table 3.4, we assume that every resident in family units has the same implicit price against the share of one-room units within an apartment. However, residents of family units could have different tastes depending on their unit's floor space (i.e., their income). Richer families may dislike one-room units more. To test this, column (2) in Table 3.5 includes the interaction term of the share of one-room units by the unit's floor space as another main regressor. The results show that the main regressors are significant in panel A but not in panel B. It means that implicit prices of one-room units are smaller for richer families in panel A but not different in panel B, which contradicts to our hypothesis. We also assume a linear relationship between

[14]This new standard for earthquake-resistant housing design was enacted in June 1981 following an amendment of the apartment standard law. This standard applies to apartments that gained their apartment certification after June 1981. Thus, some apartments built in 1982 or after could have gained their certification under the old standard.

rents and the share of one-room units within an apartment in Table 3.4. However, the effect of a 1 percentage point increase in the share of one-room units may be smaller if the majority of units within an apartment are one-room units. To test this, we include the squared share of one-room units as an additional regressor in column (3) in Table 3.5. As in the interaction term, the regressors in both panels are significantly positive. The estimated coefficients show that there is a U-shaped relationship between rents and the share of one-room units. The bottom of the "U" occurs when the share of one-room units is close to 80% (resp. 60%) in panel A (resp. panel B). Thus, when the share of one-room units smaller than 40m^2 (resp. 25m^2) is less (more) than 60% (resp. 80%) in an apartment, an increased share of one-room units decreases (increases) the rents of family units. To verify this result differently, we divide the sample into two subsamples: (i) family units in the mixed apartments that share one-room units, which are less than 50%, and (ii) family units in the mixed apartments sharing one-room units are equal to or larger than 50%. The results of the first subsample, shown in column (4) of Table 3.5, indicate that the share of one-room units is significantly negative. However, the results of the second subsample in column (5) show that the share of one-room units is insignificant. It is implied that negative externalities of one-room units occur especially when family units are majority inside apartments. Column (4) also shows that negative externalities exist even when the range of the one-room units sharing is reduced, resulting in lessening quality differences among apartments.

[Table 3.5]

To examine the effect of one-room units depending on their size, the estimations in Table 3.6 use the share of one-room units in each floor space category within an apartment as the main regressors. There are five categories of one-room units: 15–20m^2, 20–25m^2, 25–30m^2, 30–35m^2, and 35–40m^2. The specification model of each column corresponds to that in Table 3.4. For all the columns, the shares of one-room units with every category are significantly negative. As the one-room units get smaller, the absolute value of the corresponding estimated coefficient gets larger. These results imply that negative externalities are generated more from smaller one-room units.

[Table 3.6]

Overall, the estimated results show that family units in apartments with a larger share of one-room units have lower rents, indicating that negative externalities of one-room units exist similar to the regression results using the mixed floor dummy. Specifically, a 1 percentage point increase in the share of one-room units lowers the rents of family units by around 0.05% in mixed apartments. The estimation using the categorical share of one-room units also shows that one-room units with less floor space generate negative externalities.

3.5.2 Externalities within a district

We have thus far focused on the neighborhood externalities of one-room units within an apartment and found evidence of their existence. Do these externalities also affect people within the same district or just in the same apartment? Here, we analyze whether the rents of family units decline when the share of households living in one-room units rises within a district. In order to check this, we use specification model (3.3). A one-room unit is defined as one with floor space below 30m^2. Our observations for the following estimation are units of more than 40m^2 in mixed or family apartments. We include the apartment fixed effects instead of the district fixed effects to control for the unobserved time-invariant effects of each apartment.[15]

Table 3.7 presents the estimation results. As shown in columns (1)–(5), the main regressor is robustly significantly negative, indicating that the negative externalities of one-room units within a district also exist. The results indicate that a 1 percentage point increase in the district-level share of one-room residents could lower the rents of family units by around 0.086 – 0.101%. To exclude low-quality units from the sample, column (3) drops those units in densely populated urban districts and column (4) drops those built before 1982. The estimated coefficients of the main regressors in columns (3) and (4) are similar to those in column (2). Column (5) uses the share of young people in districts as an additional control. Since the majority of people living in one-room units are young people, the concentration of young people affects the demand for these types of units. The correlation coefficient between the share of one-room units and the share of young people is 0.626. The estimated coefficients in column (5) are almost same as those in column (2), which shows the robustness of the results. Column (6) includes the time-variant effects in each municipalities, which decreases the value of the estimated coefficient and the significance for the share of one-room units.

3.6 Conclusion

This study investigated the existence of the neighborhood externalities of residents of one-room units, which is a novel extension to the literature on this topic. By using rich rental data in Tokyo's 23 wards, we estimated the hedonic price equations of family units and examined whether the existence of one-room units within the same building lowers the rents of family units. The results showed that the rents of family units get decreased significantly if the one-room units exist on the same floor. A 1 percentage point increase in the share of one-room units in an apartment lowers the rents of family units in mixed apartments by about 0.05% if we define units smaller than 40m^2 as one-room units. By using the Population Census, it was also shown that a 1 percentage point increase in the share of one-room units within a district lowers the rents of family units by around 0.086 – 0.101%. In conclusion, we found clear evidence of negative neighborhood externalities both within apartments and within districts. To the best of our knowledge, this is the first study to identify neighborhood externalities within an apartment, which are the most micro-level externalities.

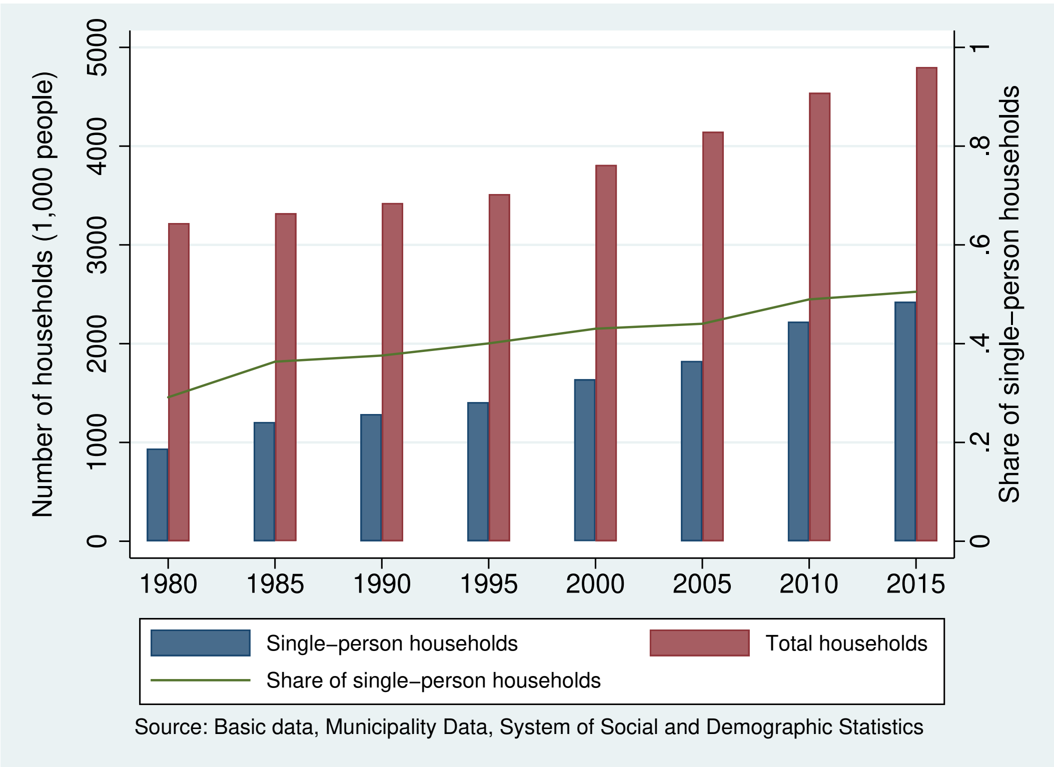
According to the results of analyses 2 and 3, neighborhood externalities of one-room units within districts are larger than the externalities within apartments. What creates the difference in volume between them? We consider that the main source of the externalities within apartments is bad-manners of one-room residents. On the other hand, the main source of the externalities within districts includes one-room residents' lack of involvement in community activities as well as their negative behavior. This lack seemingly contributes to increasing the externalities within districts. Moreover, because residents are sorted into apartments by rents, family

[15]We cannot include the unit-level fixed effects because the unit ID is not necessarily unique.

residents have more of the similar attributes with one-room residents within same apartments than one-room residents within same districts. Manners of neighbors with similar attributes are usually less problematic, and it seems to lessen the externalities within apartments compared with the externalities within districts. Identifying the causes of the externalities requires other identification strategies, but this remains for future works.

The existence of the negative externalities of one-room units suggests that the regulations on one-room apartments in Japan are a suitable policy for correcting the market failure they cause. Nevertheless, the welfare effect of the regulations requires careful discussion because demand for one-room units is increasing in Tokyo's 23 wards. The regulations may raise the rents of one-room units too far and one-room residents may then be forced to live in lower-quality units. Another direction of future work is to assess suppliers' decisions about each unit's rent and types of apartments. Since suppliers decide the rents of all the units within an apartment jointly, a cross - subsidization of rents among units may exist.[16] Although this study discussed their decision-making problems on types of apartments as carefully as possible by using district fixed effects or the time-variant additional control variable, further studies are needed to examine the effect of those regulations on rental markets and welfare.

Figure 3.1: Number of single-person households in Tokyo's 23 wards



[16] For example, under the one-room apartment regulations, suppliers have to include a certain share of family units despite their lower demand. Thus, they have an incentive to lower the rents of family units and instead raise the rents of one-room units. Another specification strategy is needed to identify the effects of cross-subsidization.

Table 3.1: Minimum and target levels of living floor space

Number of people per household	A: Minimum level	B: Target level (urban area)	Share of households in Tokyo's 23 wards (2013)		
			area < A	A ≤ area < B	B ≤ area
1	25 m^2	40 m^2	0.267	0.344	0.311
2	30 m^2	55 m^2	0.116	0.397	0.443
3	40 m^2	75 m^2	0.147	0.516	0.304
4	50 m^2	95 m^2	0.172	0.652	0.153

[1] Source: Housing and Land Survey in 2013. To be accurate, the share of households in Tokyo's 23 wards represents the share of households living in apartments in Tokyo's 23 wards.
[2] Calculation formulas for the minimum level are as follows: (1) single-person household: 25m^2; (2) multiple-person household: 10$m^2\times$ the number of people + 10m^2. Calculation formulas for the target level of the urban area are as follows: (1) single-person household: 40m^2; (2) multiple-person household: 20$m^2\times$ the number of people + 15m^2. Children of below three years, three to five years, and six to nine years are counted as 0.25 people, 0.5 people, and 0.75 people, respectively.

Table 3.2: Types of apartments

Apartment type	No.	%
Family-type apartments	37,406	24.9
Mix-type apartments	64,678	43.1
Oneroom-type apartments	47,976	32.0
Total	150,060	100.0

[1] The unit of observation is an apartment. Units with below 40m^2 are defined as one-room units.

Table 3.3: Regression results by type of apartment floor

	(1)	(2)	(3)	(4)	(5)	(6)
	A. One-room units: below 40m^2					
Mix type floor	-0.044***	-0.003***	-0.003***	-0.002***	-0.002***	-0.003***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	1020936	1020936	993540	902117	976029	1020929
R^2	0.879	0.959	0.959	0.962	0.976	0.960
adj. R^2	0.876	0.958	0.958	0.961	0.972	0.959
	B. One-room units: below 25m^2					
Mix type floor	-0.022***	-0.003***	-0.003***	-0.002**	-0.002**	-0.003***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	425224	425224	415678	380417	404770	425220
R^2	0.878	0.956	0.956	0.959	0.975	0.957
adj. R^2	0.874	0.954	0.955	0.958	0.970	0.956
Apartment F.E.	✓	✓	✓	✓	✓	✓
Year-month F.E.	✓	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Omit units in densely populated districts			✓			
Omit units built under old standards				✓		
Apartment × Year F.E.					✓	
Municipality × Year F.E.						✓

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2] In Panel A (resp. Panel B), units smaller than 40m^2 (resp. 25m^2) are defined as one-room units.

Table 3.4: Regression results by the share of one-room units within an mix type apartment

	(1)	(2)	(3)	(4)	(5)	(6)
	A. One-room units: below 40m^2					
Share of one-room units	-0.092***	-0.048***	-0.049***	-0.049***	-0.050***	-0.048***
	(0.013)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
N	319382	319368	311731	319360	317833	319441
R^2	0.677	0.913	0.913	0.916	0.935	
adj. R^2	0.675	0.912	0.912	0.916	0.933	
	B. One-room units: below 25m^2					
Share of one-room units	-0.082***	-0.048***	-0.048***	-0.048***	-0.057***	-0.057***
	(0.023)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)
N	87464	87455	86094	87441	86485	87838
R^2	0.762	0.921	0.921	0.926	0.947	
adj. R^2	0.759	0.920	0.920	0.925	0.944	
District F.E.	✓	✓	✓	✓	✓	✓
Year-month F.E.	✓	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Omit units in densely populated districts			✓			
Municipality × Year F.E.				✓		
District × Year F.E.					✓	
PDS lasso						✓

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2] In Panel A (resp. Panel B), units smaller than 40m^2 (resp. 25m^2) are defined as one-room units.

Table 3.5: Regression results by the share of one-room units within an mix type apartment with additional regressors

	(1)	(2)	(3)	(4)	(5)
	A. One-room units: below 40m^2				
Share of one-room units	-0.048***	-0.136***	-0.111***	-0.077***	0.007
	(0.007)	(0.030)	(0.029)	(0.022)	(0.021)
Share of one-room units times floor space		0.002***			
		(0.001)			
Sq. Share of one-room units			0.069**		
			(0.028)		
N	319368	319368	319368	280757	38533
R^2	0.913	0.913	0.913	0.916	0.928
adj. R^2	0.912	0.912	0.912	0.916	0.925
	B. One-room units: below 25m^2				
Share of one-room units	-0.048***	0.070	-0.162***	-0.104***	-0.027
	(0.014)	(0.054)	(0.040)	(0.029)	(0.028)
Share of one-room units times floor space		-0.003**			
		(0.001)			
Sq. Share of one-room units			0.138***		
			(0.043)		
N	87455	87455	87455	80539	6875
R^2	0.921	0.921	0.921	0.924	0.928
adj. R^2	0.920	0.920	0.920	0.923	0.922
District F.E.	✓	✓	✓	✓	✓
Year-month F.E.	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2] In Panel A (resp. Panel B), units smaller than 40m^2 (resp. 25m^2) are defined as one-room units.
[3] Column (1) in this table is the same as in column (2) in Table 3.4.

Table 3.6: Regression results by the share of one-room units within an mix type apartment in each category

	(1)	(2)	(3)	(4)	(5)	(6)
Share of one-room units (10-15m^2)	-0.419***	-0.174***	-0.173***	-0.184***	-0.131***	-0.203***
	(0.101)	(0.047)	(0.047)	(0.046)	(0.041)	(0.047)
Share of one-room units (15-20m^2)	-0.252***	-0.081**	-0.084**	-0.082**	-0.087**	-0.077**
	(0.058)	(0.033)	(0.034)	(0.034)	(0.038)	(0.033)
Share of one-room units (20-25m^2)	-0.162***	-0.064***	-0.063***	-0.063***	-0.066***	-0.071***
	(0.018)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Share of one-room units (25-30m^2)	-0.059***	-0.040***	-0.040***	-0.042***	-0.044***	-0.052***
	(0.016)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Share of one-room units (30-35m^2)	-0.083***	-0.047***	-0.048***	-0.045***	-0.046***	-0.051***
	(0.022)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Share of one-room units (35-40m^2)	-0.111***	-0.039***	-0.040***	-0.038***	-0.036**	-0.044***
	(0.029)	(0.015)	(0.015)	(0.014)	(0.016)	(0.014)
District F.E.	✓	✓	✓	✓	✓	✓
Year-month F.E.	✓	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Omit units in densely populated districts			✓			
Municipality × Year F.E.				✓		
District × Year F.E.					✓	
PDS lasso						✓
N	319382	319368	311731	319360	317833	319441
R^2	0.679	0.913	0.913	0.916	0.935	
adj. R^2	0.677	0.912	0.912	0.916	0.933	

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[2] In Panel A (resp. Panel B), units smaller than 40m^2 (resp. 25m^2) are defined as one-room units.

Table 3.7: Regression results by the share of one-room units within a district

	(1)	(2)	(3)	(4)	(5)	(6)
Share in one-room units	-0.091***	-0.087***	-0.089***	-0.101***	-0.086***	-0.014
	(0.020)	(0.018)	(0.018)	(0.020)	(0.019)	(0.018)
Apartment F.E.	✓	✓	✓	✓	✓	✓
Year-month F.E.	✓	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Omit units in densely populated districts			✓			
Omit units built under old standards				✓		
Share of young people in districts					✓	
Municipality F.E. × Year F.E.						✓
N	1367488	1367488	1333472	1206736	1367488	1367488
R^2	0.898	0.963	0.963	0.966	0.963	0.964
adj. R^2	0.894	0.961	0.962	0.964	0.961	0.962

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[2] The registered years of rentals are 2001–2005 and 2011–2015.

3.7 Appendices

Appendix A: Data construction method

Since the original data do not have the number of units or one-room units in each apartment, we construct two main regressors (the mixed dummy and the share of one-room units) by using unit ID and apartment ID. Unit ID is a number given when real estate stores construct the property data, and thus it is not unique.[17] If several real estate stores want to register the same unit in a different month, another ID will be given. However, if they do so in the same month, At Home removes duplicate units. The registered unit is released on the site for a maximum of 45 days. If the unit stays vacant, the real estate store can update the registered date and release it again with the same unit ID for up to more than 45 days. To remove these overlapping data, we sort the observations in order of date among each unit ID and calculate the date difference between the observations. If the date difference is less than 45 days, we exclude the observation of a newer date. Since the apartment ID is arranged by provider, it is unique throughout the dataset.

The first main regressor, the mixed floor dummy, takes one if at least one unit ID smaller than 40m^2 exists on the same floor. Since the non-uniqueness of unit ID is not a problem for constructing this variable, all units are used here. The second main regressor, the share of one-room units in the apartment, is calculated by counting the number of distinct unit IDs and unit IDs smaller than 40m^2 and then dividing the latter by the former. We only use newly built units to construct this variable as explained in the next paragraph. In the analyses with this second main regressor, we only use the units with the apartment ID which share of the units can be calculated.

Since unit ID is not unique as mentioned in the above, we may count the same unit multiple times. This is problematic, especially when calculating the share of one-room units, since we need to know the exact number of each type of room. To avoid this problem, newly built units are used to construct this variable. Newly built units can be specified by a variable indicating whether the unit is new. This implementation works because all the units within the apartment are considered to be registered when apartments are newly built. Moreover, even if multiple real estate agencies registered the same unit, it would be done simultaneously. Since the website provider removes duplicate units registered in the same month, we are unlikely to have multiple unit IDs for a unit. Thus, we can calculate the number of units within an apartment with few measurement errors by using only newly built units.

The data construction procedure consists of the following steps. Firstly, we remove the overlapped observations described above, which reduces the sample size to 44.7% of the original dataset. Secondly, we construct the mixed floor dummy and the share of one-room units. Thirdly, we calculate the 1st and 99th percentiles of rents and each control variable and remove units that have a value below the 1st percentile and over the 99th percentile for any variables.[18] Finally, we exclude one-room units from the dataset.

[17] Unit ID may not be unique if real estate stores create new property data when the unit is registered for advertising again after several months or years, for example. In addition, if real estate stores use a unit ID for several units mistakenly, the unit ID may not be unique.

[18] We implement this procedure for the control variables except the year and structural material.

Appendix B: Summary statistics

Table 3.8: Summary statistics of control variables

	Specification (3.1)				**Specification (3.2)**				**Specification (3.3)**	
	Panel A		Panel B		Panel A		Panel B			
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Price (log)	11.88	0.32	11.86	0.30	12.06	0.25	12.02	0.24	11.86	0.32
Mixed floor dummy ($40m^2$)	0.63	0.48								
Mixed floor dummy ($25m^2$)			0.45	0.50						
Share of OR units in apartments ($40m^2$)					0.21	0.21				
Share of OR units in apartments ($25m^2$)							0.16	0.19		
Share of OR units in districts									0.24	0.10
Floor of unit	4.99	3.54	5.10	3.57	6.46	4.19	7.00	4.13	4.50	3.32
Floor space (m^2)	51.01	10.92	50.35	10.62	50.90	10.46	48.20	8.75	54.39	12.54
Walking minutes to station					5.82	3.45	5.28	3.03		
Number of floor above ground					11.49	6.21	10.85	4.65		
Age of building (year)					4.72	3.99	5.00	3.73		
N	1024473		427012		319433		87516		1379617	

Appendix C: Regression results for the control variables

Table 3.9: Regression results for the control variables

	(1)	(2)	(3)	(4)	(5)
Floor of unit	0.008^{***}	0.008^{***}	0.008^{***}	0.009^{***}	0.008^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Floor space (m^2)	0.013^{***}	0.013^{***}	0.014^{***}	0.014^{***}	0.012^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Walking time to station			-0.002^{***}	-0.001^{**}	
			(0.001)	(0.001)	
Number of floor above ground			0.002^{***}	0.002	
			(0.001)	(0.001)	
Age of building (year)			-0.007^{***}	-0.008^{***}	
			(0.001)	(0.001)	
N	1020936	425224	319368	87455	1367488
R^2	0.959	0.956	0.913	0.921	0.963
adj. R^2	0.958	0.954	0.912	0.920	0.961

[1] Standard errors in parentheses are clustered by apartment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[2] This table only shows the estimated coefficients of the continuous control variables. The corresponding estimations for each column are as follows: (1) column (2) in Panel A of Table 3.3, (2) column (2) in Panel B of Table 3.3, (3) column (3) in Panel A of Table 3.4, (4) column (3) in Panel B of Table 3.4, (5) column (2) of Table 3.7.

References

Accetturo, Antonio, Alberto Dalmazzo, and Guido De Blasio (2014) "Skill polarization in local labor markets under share-altering technical change," *Journal of Regional Science*, Vol. 54, No. 2, pp. 249–272.

Acemoglu, Daron and Pascual Restrepo (2017) "Robots and jobs: evidence from US labor markets," NBER Working Paper No. 23285.

Ahlfeldt, Gabriel M. and Arne Feddersen (2018) "From periphery to core: measuring agglomeration effects using high-speed rail," *Journal of Economic Geography*, Vol. 18, No. 2, pp. 355–390.

Ahrens, Achim, Christian B. Hansen, and Mark E Schaffer (2018) "pdslasso and ivlasso: Programs for post-selection and post-regularization OLS or IV estimation and inference," http://ideas.repec.org/c/boc/bocode/s458459.html.

Akamatsu, Takashi, Tomoya Mori, Minoru Osawa, and Yuki Takayama (2017) "Spatial Scale of Agglomeration and Dispersion: Theoretical Foundations and Empirical Implications," RIETI Discussion Paper Series 17-E-125.

Angrist, Joshua D. and Jörn-Steffen Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010) "Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850-1860," *Social Science History*, Vol. 34, No. 2, pp. 171–197.

Audretsch, David B., Dirk Dohse, and João Pereira dos Santos (2017) "Do toll-free highways foster firm formation and employment growth? Results from a quasi-natural experiment," Kiel Working Paper No. 2080.

Autor, David H. (2015) "Why are there still so many jobs? The Hisory and Furure of Workplace Automation," *Journal of Economic Perspectives*, Vol. 29, No. 3, pp. 3–30.

Autor, David H. and David Dorn (2013) "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, Vol. 103, No. 5, pp. 1553–1597.

Autor, David H., Frank Levy, and Richard J. Murnane (2003) "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, Vol. 118, No. 4, pp. 1279–1333.

Banerjee, Abhijit V., Esther Duflo, and Nancy Qian (2012) "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Paper No.17897.

Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt (2018) "Does investment in national highways help or hurt hinterland city growth?," *Journal of Urban Economics*, Vol. 000, pp. 1–19.

Baum-Snow, Nathaniel, Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt (2016) "Highways, Market Access and Urban Growth in China," SERC Discussion Papers SERCDP0200.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa (2008) "Place of work and place of residence: Informal hiring networks and labor market outcomes," *Journal of Political Economy*, Vol. 116, No. 6, pp. 1150–1196.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014) "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, Vol. 81, No. 2, pp. 608–650.

Berger, Thor and Kerstin Enflo (2017) "Locomotives of local growth: The short- and long-term impact of railroads in Sweden," *Journal of Urban Economics*, Vol. 98, pp. 124–138.

Berger, Thor and Carl Benedikt Frey (2016) "Did the Computer Revolution shift the fortunes of U.S. cities? Technology shocks and the geography of new jobs," *Regional Science and Urban Economics*, Vol. 57, pp. 38–45.

Berman, Eli (2000) "Does Factor-Biased Technological Change Stifle International Convergence? Evidence from Manufacturing," NBER Working Paper No. 7964.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito (2019) "Production Networks, Geography, and Firm Performance," *Journal of Political Economy*, Vol. 127, No. 2, pp. 639–688.

Boyle, Melissa A., Katherine A. Kiel, Melissa A. Boyle, and Katherine A. Kiel (2001) "A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities," *Journal of Real Estate Literature*, Vol. 9, No. 2, pp. 117–144.

Chandra, Amitabh and Eric Thompson (2000) "Does public infrastructure affect economic activity? Evidence from the rural interstate highway system," *Regional Science and Urban Economics*, Vol. 30, No. 4, pp. 457–490.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor (2008) "When are ghettos bad ? Lessons from immigrant segregation in the United States," *Journal of Urban Economics*, Vol. 63, pp. 759–774.

Datta, Saugato (2012) "The impact of improved highways on Indian firms," *Journal of Development Economics*, Vol. 99, No. 1, pp. 46–57.

Davies, Paul S., Michael J. Greenwood, and Haizheng Li (2001) "A Conditional Logit Approach to U.S. State-to-State Migration," *Journal of Regional Science*, Vol. 41, No. 2, pp. 337–360.

Diao, Mi, Yu Qin, and Tien Foo Sing (2016) "Negative Externalities of Rail Noise and Housing Values: Evidence from the Cessation of Railway Operations in Singapore," *Real Estate Economics*, Vol. 44, No. 4, pp. 878–917.

Donaldson, Dave (2018) "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *The American Economic Review*, Vol. 108, No. 4-5, pp. 899–934.

Donaldson, Dave and Richard Hornbeck (2016) "Railroads and American Economic Growth: A "Market Access" Approach," *The Quarterly Journal of Economics*, Vol. 131, No. 2, pp. 799–858.

Duranton, Gilles, Peter M. Morrow, and Matthew A. Turner (2013) "Roads and trade: Evidence from the US," *Review of Economic Studies*, Vol. 81, No. 2, pp. 681–724.

Duranton, Gilles and Matthew A. Turner (2012) "Urban Growth and Transportation," *Review of Economic Studies*, Vol. 79, No. 4, pp. 1407–1440.

Executive Office of the President of the United States of America (2016) "Artificial Intelligence, Automation, and the Economy,"Technical report, Washington, DC.

Ezawa, Akira and Yoshihide Nakagawa (2009) "The comparison of demands and stocks for studio apartments in Tokyo's 23wards (In Japanese)," *Proceedings of Infrastructure Planning (CD-ROM)*, Vol. 40, p. ROMBUNNO.27.

Faber, Benjamin (2014) "Trade integration, market size, and industrialization: Evidence from China's national trunk highway system," *Review of Economic Studies*, Vol. 81, No. 3, pp. 1046–1070.

Frey, Carl Benedikt and Michael A. Osborne (2017) "The future of employment: How susceptible are jobs to computerisation?," *Technological Forecasting and Social Change*, Vol. 114, pp. 254–280.

Fu, Yuming and Stuart A. Gabriel (2012) "Labor migration, human capital agglomeration and regional development in China," *Regional Science and Urban Economics*, Vol. 42, No. 3, pp. 473–484.

Fujita, M. (1989) *Urban economic theory: land use and city size*, Cambridge: Cambridge University Press.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (2001) *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.

Fujita, Masahisa and Jacques François Thisse (2013) *Economics of agglomeration: cities, industrial location, and globalization*, New York, NY: Cambridge university press, 2nd edition.

Gaspar, Jess and Edward L. Glaeser (1998) "Technology and the future of cities," *Journal of Urban Economics*, Vol. 43, pp. 136–156.

Gerardi, Kristopher, Eric Rosenblatt, Paul S. Willen, and Vincent Yao (2015) "Foreclosure externalities: New evidence," *Journal of Urban Economics*, Vol. 87, pp. 42–56.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr (2016) "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing," *Economic Journal*, Vol. 126, No. 591, pp. 317–357.

Gibbons, Stephen, Teemu Lyytikäinen, Henry G. Overman, and Rosa Sanchis-Guarner (2019) "New road infrastructure: The effects on firms," *Journal of Urban Economics*, Vol. 110, pp. 35–50.

Goos, Maarten and Alan Manning (2007) "Lousy and Lovely Jobs: the rising polarization of work britain," *Review of Economics and Statistics*, Vol. 89, No. 1, pp. 118–133.

Greenwood, Michael J. (1997) "Chapter 12 Internal migration in developed countries," *Handbook of Population and Family Economics*, Vol. 1, No. PART B, pp. 647–720.

Greenwood, Michael J. and Gary L. Hunt (1989) "Jobs versus amenities in the analysis of metropolitan migration," *Journal of Urban Economics*, Vol. 25, No. 1, pp. 1–16.

Handel, Michael J. (2016) "The O*NET content model: strengths and limitationsStärken und Grenzen des O*NET-Models," *Journal for Labour Market Research*, Vol. 49, No. 2, pp. 157–176.

Helpman, Elhanan (1998) "The size of regions," in Pines, David, Efraim Sadka, and Itzahk Zilcha eds. *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge, MA: Cambridge University Press, pp. 33–54.

Heurermann, Daniel F. and Johannes F. Schmieder (2014) "Warping Space: High-Speed Rail and Returns to Scale in Local Labor Markets," Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Local Labour Markets, No. A13-V1.

Hornung, Erik (2015) "Railroads and growth in Prussia," *Journal of the European Economic Association*, Vol. 13, No. 4, pp. 699–736.

Ioannides, Yannis M. and Jeffrey E. Zabel (2008) "Interactions, neighborhood selection and housing demand," *Journal of Urban Economics*, Vol. 63, No. 1, pp. 229–252.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani (2013) *An introduction to Statistical Learning*, New York, NY: Springer.

Jedweb, Remi and Alexander Moradi (2016) "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa," *The Review of Economics and Statistics*, Vol. 98, No. 2, pp. 268–284.

Kanemoto, Yoshitsugu (1980) *Theories of Urban Externalities*, Amsterdam: North Holland Publishing Co.

Kanemoto, Yoshitsugu and Kazuyuki Tokuoka (2002) "Proposal for the Standards of Metropolitan Areas of Japan," *Journal of Applied Regional Science*, Vol. 7, pp. 1–16.

Kawaguchi, Daiji and Norifumi Yukutake (2017) "Estimating the residential land damage of the Fukushima nuclear accident," *Journal of Urban Economics*, Vol. 99, pp. 148–160.

Kinoshita, Ryuji, Toshio Otsuki, and Kahori Fukami (2008) "A study on one-room apartments problems and the relevant policies in 23 wards of Tokyo (In Japanese)," *Journal of Architecture and Planning*, Vol. 73, No. 624, pp. 263–270.

Krugman, Paul (1991) "Increasing Returns and Economic Geography," *Journal of Political Economy*, Vol. 99, No. 3, pp. 483–499.

Kuminoff, Nicolai V. and Jaren C. Pope (2014) "Do "Capitalization Effects" for Public Goods Reveal the Public's Willingness to Pay ?," *International Economic Review*, Vol. 55, No. 4, pp. 1227–1250.

Leonard, Tammy, Nikhil Jha, and Lei Zhang (2017) "Neighborhood price externalities of foreclosure rehabilitation: an examination of the Neighborhood Stabilization Program," *Empirical Economics*, Vol. 52, No. 3, pp. 955–975.

Leung, Tin Cheuk and Kwok Ping Tsang (2012) "Love thy neighbor: Income distribution and housing preferences," *Journal of Housing Economics*, Vol. 21, No. 4, pp. 322–335.

Lin, Yatang (2017) "Travel costs and urban specialization patterns: Evidence from China's high speed railway system," *Journal of Urban Economics*, Vol. 98, pp. 98–123.

Liu, Crocker H., Stuart S. Rosenthal, and William C. Strange (2018) "The vertical city: Rent gradients, spatial structure, and agglomeration economies," *Journal of Urban Economics*, Vol. 106, pp. 101–122.

McFadden, Daniel (1973) "Conditional logit analysis of qualitative choice behavior," in Zarembka, Paul ed. *Frontiers in Econometrics*, New York: Academic Press.

Melitz, Marc J. (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, Vol. 71, No. 6, pp. 1695–1725.

Michaels, Guy (2008) "The effect of trade on the demand for skill: Evidence from the interstate highway system," *Review of Economics and Statistics*, Vol. 90, No. 4, pp. 683–701.

Michaels, Guy, Ashwini Natraj, and John Van Reenen (2014) "Has ICT Polarized Skill Demand? Evidence from eleven Countries over twenty-five Years," *The Review of Economics and Statistics*, Vol. 96, No. 1, pp. 60–77.

Ministry of Land, Infrastructure, Transport and Tourism, "About the shinkansen railway (shinkansen tetsudo ni tsuite)."

Modestino, Alicia Sasser and Julia Dennett (2013) "Are American homeowners locked into their houses? The impact of housing market conditions on state-to-state migration," *Regional Science and Urban Economics*, Vol. 43, No. 2, pp. 322–337.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak (2011) "Internal migration in the United States," *Journal of Economic Perspectives*, Vol. 25, No. 3, pp. 173–196.

Mori, Tomoya and Kohei Takeda (2018) "Highways, high-speed railways, and urban growth: evidence from Japan 1970-2014," Mimeo.

Pingle, Jonathan F. (2007) "A note on measuring internal migration in the United States," *Economics Letters*, Vol. 94, No. 1, pp. 38–42.

Plantinga, Andrew J., Cécile Détang-Dessendre, Gary L. Hunt, and Virginie Piguet (2013) "Housing prices and inter-urban migration," *Regional Science and Urban Economics*, Vol. 43, No. 2, pp. 296–306.

Poncet, Sandra (2006) "Provincial migration dynamics in China: Borders, costs and economic motivations," *Regional Science and Urban Economics*, Vol. 36, No. 3, pp. 385–398.

Qin, Yu (2017) "'No county left behind?' The distributional impact of high-speed rail upgrades in China," *Journal of Economic Geography*, Vol. 17, No. 3, pp. 489–520.

Redding, Stephen J. and Matthew A. Turner (2015) "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, Vol. 5: Elsevier B.V. pp. 1339–1398.

Rosen, Sherwin (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, Vol. 82, No. 1, pp. 34–55.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III (2010) "Housing Externalities," *The Journal of Poitical Economy*, Vol. 118, No. 3, pp. 1–8.

Saito, Tokuhiko and Hiroki Yoshihara (2008) "Regulation is needed or not: One-room apartments and Late night operation of convenience stores (Kisei wa hitsuyo-ka)," *Asahi Shimbun*.

Sanchis-Guarner, Rosa (2013) "Driving up Wages: The Effects of Road Construction in Great Britain," SERC Discussion Papers SERCDP0120.

Sandler, Danielle H. (2017) "Externalities of public housing: The effect of public housing demolitions on local crime," *Regional Science and Urban Economics*, Vol. 62, pp. 24–35.

Sasser, Alicia C. (2010) "Voting with their feet: Relative economic conditions and state migration patterns," *Regional Science and Urban Economics*, Vol. 40, No. 2-3, pp. 122–135.

Sato, Yasuhiro (2014) *Invitation letter to urban and regional economics (Toshi chiiki keizai-gaku e no shotaijo)*, Tokyo: Yuhikaku.

Shimizu, Chihiro and Masayuki Nakagawa (2018) "Mansion no Rokyu-ka to jinko no korei-ka ga motarasu toshi no sugata [The effect of deteriorated mansions and aging societies on the metropolitan areas (in Japanese)]," in Saito, Makoto ed. *Toshi no Oi [Age of the city]*, Tokyo: Keiso Shobo.

Solow, Robert M. (1957) "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, Vol. 39, No. 3, pp. 312–320.

Storeygard, Adam (2016) "Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa," *Review of Economic Studies*, Vol. 83, No. 3, pp. 1263–1295.

Tabuchi, Takatoshi (1998) "Urban Agglomeration and Dispersion : A Synthesis of Alonso and Krugman," *Journal of Urban Economics*, Vol. 44, pp. 333–351.

Tabuchi, Takatoshi, Jacques François Thisse, and Xiwei Zhu (2018) "Does Technological Progress Magnify Regional Disparities?," *International Economic Review*, Vol. 59, No. 2, pp. 647–663.

Tojo, Misao (1927) *A map of the dialects of Japan [Dainihon hogen chizu]*, Tokyo: Ikuei Shoin.

Tonozuka, Yasushi (2010) "Wanrumumanshon kenchiku kisei no mondaiten to kaihatsu jigyosha no torikumi [The issues on one-room apartment regulations and actions of developers (in Japanese)]," *Toshi Jutaku Gaku [Urban Housing Studies]*, Vol. 2010, No. 70, pp. 37–41.

United Nations, Department of Economic and Social Affairs, Population Division (2018) "World Urbanization Prospects: The 2018 Revision," Online Edition.

U.S. Census Bureau (2018) "Understanding and using American Community Survey data: What all data users need to know,"Technical report, Washington, DC.

Violante, Giovanni L. (2008) "Skill-Biased Technical Change," *The New Palgrave Dictionary of Economics*, pp. 520–523.

Volpe Martincus, Christian and Juan Blyde (2013) "Shaky roads and trembling exports: Assessing the trade effects of domestic infrastructure using a natural experiment," *Journal of International Economics*, Vol. 90, No. 1, pp. 148–161.

West Japan Railway Company (2011) "About timetable revision in the spring of 2011 (heisei 23-nen haru daiya kaisei ni tsuite),"Technical report.

Winke, Tim (2016) "The impact of aircraft noise on apartment prices: a differences-in-differences hedonic approach for Frankfurt, Germany," *Journal of Economic Geography*, Vol. 17, No. December 2016, p. lbw040.

Zabel, Jeffrey E. (2012) "Migration, housing market, and labor market responses to employment shocks," *Journal of Urban Economics*, Vol. 72, No. 2-3, pp. 267–284.

Zhang, Lei and Tammy Leonard (2014) "Neighborhood impact of foreclosure: A quantile regression approach," *Regional Science and Urban Economics*, Vol. 48, pp. 133–143.