Commuting Zones in Japan*

Daisuke Adachi† Taiyo Fukai‡ Daiji Kawaguchi§ Yukiko U. Saito¶

April 30, 2025

Abstract

Choosing a proper geographic unit is crucial for an empirical analysis of local labor markets. To define appropriate local labor markets for Japan, we first construct commuting zones (CZs) using the commuting patterns observed in the Population Census from 1980 to 2020 and the hierarchical agglomerative clustering method adopted by Tolbert and Sizer (1996) to delineate CZs in the US. From 1,736 municipalities in 2015, for example, we constructed 265 CZs that are mutually exclusive and exhaustive. We then compared the properties of CZs with those of other potential administrative units including the municipality, prefecture and Urban Employment Area (UEA) proposed by Kanemoto and Tokuoka (2002), finding that our proposed CZs capture the actual commuting patterns and heterogeneity of local labor markets.

Keywords: Commuting zones, Regional economies, Hierarchical agglomerative clustering, Geographic boundaries.

JEL Classification: J61, R12.

---

*This study was conducted under the "Dynamics of Inter-organizational Network and Firm Life Cycle" project of the Japan Research Institute of Economy, Trade and Industry (RIETI) and financially supported by grants from the Japan Science and Technology Agency's Research Institute of Science and Technology for Society (JPMJRX18H3) and from the Japan Society for the Promotion of Science (KAKENHI Grant Number 15H05692; PI: Hidehiko Ichimura). The authors are grateful for helpful comments and suggestions by Keisuke Kawata and Discussion Paper seminar participants at RIETI. We also thank Shuhei Kainuma for outstanding research assistance and Philip MacLellan for editing assistance. All remaining errors are our own. Correspondence tables of municipality code and CZ code for 1980 to 2020 and all replication files are already published and available at `https://github.com/daisukeadachi/commuting_zone_japan`. Please cite this paper if you use any of these results.

†Aarhus University and RIETI. E-mail: daisuke.adachi.econ@au.dk

‡Gakushuin University. E-mail: taiyo.fukai@gakushuin.ac.jp

§University of Tokyo, RIETI and IZA. E-mail: kawaguchi@e.u-tokyo.ac.jp

¶Waseda University and RIETI. E-mail: saito-yukiko@waseda.jp

1 Introduction

Empirical economists have long used geographic units as the unit of analysis, and the growing popularity of the shift-share method introduced by Bartik (1991) and Altonji and Card (1991) has further increased the importance of local labor markets in properly understanding a wide range of economic issues, particularly labor market outcomes. Specifically, since the shift-share method aims to exploit the variation in local characteristics such as industrial and demographic composition, it is critical to correctly choose the relevant unit for each analysis purpose. Instead of traditional administrative units such as states and counties, many researchers in the US have begun to use commuting zones (CZs), proposed by Tolbert and Killian (1987) and Tolbert and Sizer (1996).[1] This is because CZs reflect integrated labor markets better than administrative units due to commuting across administrative units. However, to the best of our knowledge, there has been no attempt to date to delineate CZs in Japan. This study thus attempts to fill the gap by first developing a set of CZs for Japan, using the Tolbert and Sizer (1996) hierarchical agglomerative clustering (HAC) method and micro data from the 1980-2020 quinquennial Population Census indicating the municipalities in which each individual lives and works. After deriving the CZs, we assess their performance by comparing them with administrative geographic units.

The contribution of our paper is delineating commuting zones in Japan based on a standard method, paying careful attention to the proper choice of the threshold parameter in the Japanese context. The comparison of CZs with the Urban Employment Area (UEA) proposed by Kanemoto and Tokuoka (2002) clarifies the benefit of delineating the CZs. While CZs and UEAs are similar in that they both consider cross-jurisdictional commuting patterns, the UEA, as the name suggests, is exclusively focused on areas around urban city centers as defined by the densely inhabited district. Because the merging process begins from the urban centers, many rural municipalities end up not belonging to a UEA. Thus UEAs do not comprise an exhaustive set of geographic units within a country. CZs, by contrast, are not defined by urban city centers and so all administrative units belong to a CZ regardless of how small or remote. CZs thus comprise a set of mutually exclusive and exhaustive geographic units, which helps to increase geographic coverage. This increased coverage especially helps researchers when their scope of analysis is not restricted to urban city centers but also includes rural districts such as the impact of factory opening on the local economy (Greenstone et al., 2010). Section 2 discusses the differences between UEA and CZ in detail.

This paper first delineates the CZs in Japan using the 1980-2020 waves of the quinquennial

[1]See, for example, Autor et al. (2013); Chetty et al. (2014); Acemoglu and Restrepo (2020).

Population Census, which is conducted based on residence but also records work location. To construct the CZs, we began with municipalities (*Shi Ku Cho Son*) as the smallest geographic unit and then progressively merged them using the Tolbert and Sizer's (1996) hierarchical agglomerative clustering (HAC) method, based on the frequency of commuting between municipalities compared to the size of their work forces. For 2015, this merging process yielded 265 mutually exclusive and exhaustive CZs from the 1736 municipalities in Japan's 47 prefectures.

The HAC method is a standard statistical clustering method in the machine learning literature. Specifically, it begins by initially treating each municipality as a single geographic unit, with the "distance" between any two municipalities defined by the number of commuters relative to the number of workers in the two municipalities. The larger the proportion of commuters relative to workers, the shorter the distance. Two municipalities are merged to create a group if the distance between the two is smaller than a cutoff level (explained below). At this stage, a group is either a single municipality or a set of municipalities. In the second stage, the distance between two groups is defined as the average of the distances between all possible combinations of municipalities belonging to the two groups, and two groups are merged if the distance between them is smaller than the cutoff. This process is iterated until all the distances between clusters are larger than the cutoff. The groups remaining at the end of the iteration is defined as CZs. As is clear from this procedure, the choice of the cutoff for the distance metric is crucial. We follow Tolbert and Sizer (1996) in setting the cutoff because the distance metric, the frequency of commuting between municipalities relative to their work forces, is a standardized measurement that is free from any difference in the total population or land size across municipalities and across countries.

We find that a CZ is typically contained within a prefecture, which justifies the use of the prefecture as a analysis unit as an approximation to the true delineation for regional economic activities. However, there are several important exceptions in large metropolitan areas such as Tokyo and Osaka, whereby a single CZ includes municipalities belonging to multiple prefectures. On the other hand, we also find cases in which a prefecture includes multiple CZs. Together, these features suggest that our CZs perform better than administrative units such as prefectures when analyzing local labor markets and issues of urban economics such as agglomeration.

To test this point, we compare the performance of the proposed CZs with municipal and prefectural administrative units. First, we considered the catchment rate, the fraction of workers commuting within a geographic unit. As a result, we find that the CZs performed significantly better than municipalities but only slightly worse than prefectures. Thus, the CZs succinctly capture the actual local labor market as far as commuting pattern is concerned. In contrast, the downside of using the relatively coarse CZs instead of the more fine-grained municipality is the increased heterogeneity within the geographic unit. To investigate this point further, we examined how much additional variation in employment and wages could be explained by adding municipality fixed effects in addition to the CZ fixed effects of the main analysis. The results show that while the municipality fixed effects explain the variation in employment and wages conditional on the CZ fixed effects in statistically significant ways, the magnitude of the increase is small. Overall, we conclude that, at the very least, using the CZ instead of the prefecture as the unit of analysis increases the number of observations without sacrificing the catchment rate or homogeneity of labor market outcomes.

Lastly, we study the impact of our choice of the cutoff level of the distance between geographic units for merging in the HAC method. We find that setting a low cutoff increases the number of CZs and the homogeneity of employment probability within a CZ but decreases the catchment rate, or the probability that the commuting destination is actually contained in the residential CZ. In other words, increasing the number of CZs and improving within-CZ homogeneity comes at the cost of failing to capture actual commuting patterns, which shows the trade offs among these three variables when choosing the cutoff. We argue that our choice of the cutoff point, which follows Tolbert and Sizer (1996), is reasonable both because their derivation of CZs is widely accepted by empirical economists and because their distance metric is normalized so that it is not affected by any differences in population or land area between countries.

The paper proceeds as follows: Section 2 discusses the existing concepts of geographic units for empirical analysis and Section 3 describes the HAC method and its application to Japan's Population Census. Section 4 shows the clustering results and performance assessment, Section 5 discusses the choice of a key hyperparameter of the HAC method, and Section 6 concludes the paper.

2 A Review of Geographic Units for Economic Analysis

In addition to administrative units such as prefecture/state and municipality, several non-administrative geographic units have been used for economic analysis. This section compares these other alternatives, particularly the UEA by Kanemoto and Tokuoka (2002), to the commuting zone proposed in this paper. The UEA consists of Metropolitan Employment Areas (MEAs) whose central cities have Densely Inhabited District (DID) populations exceeding 50,000 and Micropolitan Employment Areas (McEAs) with DID populations between 10,000 and 50,000. In 2000, MEAs covered more than 80 percent of the population of Japan, (Fujita et al., 2004) and in 2015, there were 187 UEAs and 94 MEAs in the country. Kanemoto and Tokuoka (2002) defines a MEA or a McEA as an urban center and then any surrounding municipality with more than 10 percent of its residents commuting to an urban center is merged with it. As the merging process is initiated from urban centers, many rural municipalities do not belong to any UEA. Thus, UEAs consist of a mutually exclusive but non-exhaustive set of geographic units. CZs and UEAs are similar in that they both are derived from commuting patterns observed in the Population Census but, as the name suggests, UEA constructs commuting zones exclusively around urban city centers.

The US counterpart of the UEA is the Metropolitan Statistical Area (MSA), which also leaves out rural areas, and so for completeness, recent studies in the US have begun to employ CZs instead of MSAs (Autor et al., 2013; Chetty et al., 2014; Acemoglu and Restrepo, 2020). Excluding rural municipalities from the unit of analysis is potentially problematic for analyzing the effect of international trade or factor-biased technological changes (e.g. industrial robots) on manufacturing employment because some manufacturing agglomerations are located in rural company towns (Dorn, 2009). Neglecting this industrial agglomeration may thus bias the empirical results.

A notable difference between CZs and UEAs is that CZs tend to divide metropolitan areas into smaller sections, whereas UEAs are more aggregated. This discrepancy arises because the hierarchical agglomerative clustering (HAC) with average linkage used to derive CZs tends to separate two populous geographic units or a combination of one populous area and several sparsely populated ones, as discussed in Section 3.4. Since such combinations are common within metropolitan areas, CZs are more likely to split them. Nonetheless, despite their methodological differences, CZs and UEAs show remarkably similar spatial patterns, suggesting both approaches effectively capture fundamental aspects of commuting areas.

3 Method

In this section, we first briefly review existing methods for delineating local labor markets, then provide a general description of the hierarchical agglomerative clustering (HAC) method by which we construct our commuting zones (CZs). We then discuss how to apply the method to our setting, merging of regional units into clusters using the quinquennial Population Census. Finally we discuss the strengths and weaknesses of the method.

3.1 Overview

Although the concept of delineating local labor markets is simple, its implementation is difficult in practice, as one must choose between many different possible data sets and methods. First, data that captures the journey-to-work experience as we have obtained from the Population Census is reasonable to use when workers commute between regions, as this data contains dweller-commuter pair observations that provide information on worker movement between regions and thus helps to more accurately describe local labor markets. Delineation of labor markets based on this type of data is commonly used in Germany, the Netherlands, Spain, the UK, and the US (Kropp and Schwengler, 2017).

Beyond data attributes, there are additional methodological choices that have led to a lack of consensus in delineating local labor markets, particularly in the choice of thresholds. The threshold method (Kropp and Schwengler, 2017), for example, requires the specification of more than one arbitrary threshold value for commuting flows upon which two regions are separated. This paper adopts the hierarchical agglomerative clustering (HAC) method employed by Tolbert and Killian (1987) and Tolbert and Sizer (1996) in the US and Cörvers et al. (2009) in the Netherlands, as it provides a consistent and non-arbitrary method for choosing a threshold value. The HAC method is described next.

3.2 Hierarchical Agglomerative Clustering

Here we describe the process of delineating commuter zones by incrementally clustering municipalities based on their commuting distances. First consider that there is a set of nodes (initially municipalities; later clusters) with "dissimilarities" (commuting distance, formally defined below) between each pair of nodes. The problem is how to aggregate nodes into a set of clusters whereby elements are similar within a cluster but are dissimilar across clusters. HAC is a prevalent method for doing this, particularly by Tolbert and Sizer (1996) in their derivation of CZs in the US which we have followed in our delineation of CZs in Japan. We describe the HAC algorithm below.

Let the set of nodes, or groups containing a single node, be $\{1, \ldots, N\} \equiv \mathcal{N}^N$. Note that the superscript N represents the count of groups in $\mathcal{N}^N$. For any $i, j \in \mathcal{N}^N$, suppose we have a dissimilarity measure D_{ij}. The matrix composed of such dissimilarity measures is "dissimilarity matrix" $D = \left(D_{ij}\right)_{ij}$. The algorithm proceeds as follows. First, from each pair of elements, we find $(i, j) = \arg\min_{(i',j')} D_{i'j'}$. Then we aggregate i and j to create a group $c = \{i, j\}$. With a slight abuse of notation, we write the resulting set of groups as $\mathcal{N}^{N-1} \equiv \{1, \ldots, i-1, i+1, \ldots, j-1, j+1, \ldots, N, c\}$. Note again that the superscript $N-1$ represents the count of groups in $\mathcal{N}^{N-1}$. We may now define a new system of dissimilarities given D as follows: between any two groups c and d, the distance D_{cd} can be defined by the average of distances between any pair of nodes in groups c and d as follows: $D_{cd} \equiv \frac{1}{|c| \times |d|} \sum_{i \in c, j \in d} D_{ij}$. From Tolbert and Sizer (1996), this method of defining the distance is called "average linkage". Given this definition of dissimilarities between groups, we may iterate the same process of finding the nearest pair of groups and aggregate it into a group. This process continues, and at each step k of finding the nearest pair and aggregating it, the number of groups $|\mathcal{N}^{N-k}|$ declines by one. Hence, with $N-1$ steps, we arrive at the set of groups that consists of only the one that contains all the nodes, $\mathcal{N}^1 = \{\{1, \ldots, N\}\}$. Note that each step can be characterized by the largest value of aggregation D_{cd} such that groups c and d are aggregated. Moreover, it is trivial to show that such a characterizing value of each step of aggregation is increasing as each step proceeds because of the property of the minimizing problem and average linkage. Therefore, to finally define the clusters, we terminate the aggregation process at a pre-specified hyperparameter of dissimilarity cutoff h. In other words, in the final cluster, c and d are in the same group if and only if $D_{cd} < h$.

In order to visualize the HAC aggregation process with average linkage, a tree diagram called a dendrogram is useful. One can always arrange the order of elements to show a graph that represents at what value of dissimilarities the groups are formed with which pairs, and a dendrogram is created by placing the value of dissimilarities on the vertical axis with large values on the top. Examples of such trees are shown in Figures 6 and 7 of Section 5. The dendrogram provides the intuition behind particular groupings resulting from our algorithm. In such a tree diagram, h tells us at what height the tree is "cropped" and the resulting clusters should be defined, which is sometimes called the clustering "tree height". Choice of this tree height is a non-trivial problem without a consensus as yet. Discussion about the choice of this hyperparameter is discussed in Section 5.

3.3 Application of HAC to Commuting Flows

Next, following Tolbert and Sizer (1996), we apply the HAC method to the data to construct the clusters. With the idea that distance can be represented not by geographical distance but instead by commuting flows between two local units, Tolbert and Sizer (1996) use 1990 US journey-to-work data that contains for each worker, their residence county (or county-equivalent in some parts of the US) i and workplace county j. In 1990, there were 3,141 counties in the US. The definition of the dissimilarity values is the following: for any i, $D_{ii} = 0$. For $i \neq j$,

$$D_{ij} \equiv \max\left\{\epsilon, 1 - \frac{f_{ij} + f_{ji}}{\min\left(p_i, p_j\right)}\right\} \in [0, 1],$$

where f_{ij} is the number of commuters from residence i to workplace j, p_i is population of i, and ϵ is an arbitrarily small value.[2] A benefit of this definition is that the dissimilarity values lie between zero and one, so $D_{ij} \in [0, 1] \Rightarrow D_{cd} \in [0, 1]$ for any clusters c, d since $f_{ij} < p_i$ for any i. Therefore, two groups are close in terms of commuting flows if D_{cd} is small and close to zero, while they are far apart if D_{cd} is large and close to one. Finally, Tolbert and Sizer (1996) set h to be 0.98.

The choice of the dissimilarity cutoff value (0.98) deserves careful consideration. While this value follows Tolbert and Sizer (1996), we recognize that establishing a theoretically optimal cutoff is inherently challenging. To address this difficulty, we conduct a comprehensive analysis of CZ configurations across various cutoff values from 0.8 to 1.0. This analysis provides a thorough understanding of how the threshold choice affects commuting zone formation. For research transparency, we make available the CZ delineations for all cutoff values, allowing researchers to select the most appropriate threshold for their specific needs, rather than being constrained by our baseline choice. Our value of 0.98 maintains international comparability while acknowledging the inherent arbitrariness in selecting any single threshold.

To implement this method in Japan, we first obtained micro data from the 1980-2020 waves of the *Population Census* administered by the Ministry of Internal Affairs and Communications (MIC). The Census takes place in October every five years and covers all individuals in Japan, including foreign residents. It collects basic demographic information such as age, gender, marital status, employment status and work location. Employment status is determined by the actual employment/unemployment situation during the last week of September.[3] The respondent chooses either of (1) mostly worked; (2) worked besides doing housework; (3) worked besides attending school; (4) absent from work; (5) unemployed; (6) did housework; (7) attended school; (8) other (elderly persons, etc) .

[2]Following Tolbert and Sizer (1996), we set $\epsilon = 0.001$. The choice of this variable does not materially affect the clustering result.

[3]The Census defines employment status by the "actual" method, and not the "usual" method employed by some other large-scale surveys such as the ESS.

Table 1 summarizes the sample used to construct the CZs. Row (1) reports the total number of people included in each census. From there, the analysis sample was constructed by choosing the observations with valid household type, employment status and employment location. This did not eliminate many observations, as the percentage with non-missing variables was high at more than 97 percent in the 1980-2005 Population Censuses. In 2010 and 2015, the number of missing variables increased, but still covered more than 90 percent of the total population. To construct the actual CZs, we further limited the sample to households with non-institutionalized household heads.[4] Among the non-institutionalized population, we selected those household members who mainly worked. Row (6) reports the analysis sample size used to construct the CZs.

Next, using data on individual employment status and the municipalities of residence and work, we aggregated the data using the HAC method to delineate the CZs. We began the process with municipalities[5] instead of Census tracts, which are finer geographic units, because most relevant data for our purposes such as residence and workplace was available only at the municipality level. Using this information, we constructed frequency matrices of commuting flows, again restricting our sample only to those who mostly worked. Column (2) of Table 2 shows the number of municipalities in Japan during each Census year. We then applied HAC using the average linkage and dissimilarity measure defined in equation (3.3) and, following Tolbert and Sizer (1996), $h = 0.98$.

3.4 Discussion of the CZ Derivation Method

This section highlights the strengths of deriving CZs based on Tolbert and Sizer (1996)'s method in Japan as compared to the implementations in the US. In addition, this section discusses the strength and weakness of Tolbert and Sizer (1996)'s method compared to the alternative methods for delineating local labor markets.

There are several advantages of our dataset relative to those used for deriving CZs in the US. First among these is that Japan's Population Census provides more frequent, comprehensive and consistent data than is available in the US. With Japan's Population Census conducted every five years instead of every ten years in the US, the evolution of CZs can be studied on a more frequent basis in Japan. In addition, Japan's Population Census has included a question about the

[4]Non-institutionalized households are defined as those whose head of household is not a hospital inpatient, a prison inmate, an active Self-Defence Forces member, an inmate of a reformatory institution, or other social institutions.

[5]We treated municipalities with multiple administrative districts, called government-designated cities, as a single geographic unit. Government-designated cities include large cities such as Sapporo, Sendai, Tokyo, Yokohama, Nagoya, Kyoto, Osaka and Fukuoka.

Table 1: Sample Construction

	Census year		
	1980	1985	1990
Data availability			
(1) Entire sample	117,075,477	121,063,836	123,693,359
(2) Household type available	117,052,840	121,007,577	123,287,027
(3) Work status available	116,836,846	120,830,526	122,867,728
(4) Work place available	116,836,846	120,830,526	122,867,728
(Fraction of non-missing, %)	(99.8)	(99.8)	(98.2)
Sample restriction			
(5) Non-institutionalized population	115,649,464	119,489,171	121,462,573
(6) Mainly worked	46,497,979	48,269,232	52,072,091
	Census year		
	1995	2000	2005
Data availability			
(1) Entire sample	125,674,544	126,925,843	127,767,994
(2) Household type available	125,439,937	126,697,282	127,285,653
(3) Work status available	124,912,814	124,956,611	123,928,692
(4) Work place available	124,912,814	124,956,611	123,928,692
(Fraction of non-missing, %)	(98.2)	(98.4)	(97.0)
Sample restriction			
(5) Non-institutionalized population	123,437,206	123,256,450	121,885,156
(6) Mainly worked	53,595,949	53,173,075	50,808,827
	Census year		
	2010	2015	2020
Data availability			
(1) Entire sample	128,057,352	127,094,745	126,146,099
(2) Household type available	128,057,352	127,094,745	126,146,099
(3) Work status available	120,874,438	118,405,660	111,459,638
(4) Work place available	117,388,130	114,840,345	108,215,246
(Fraction of non-missing, %)	(91.7)	(90.4)	(85.8)
Sample restriction			
(5) Non-institutionalized population	115,129,468	112,299,559	105,483,376
(6) Mainly worked	47,059,134	46,556,082	46,500,277

Note: This table shows the Population Census data used to calculate the CZs. Here, non-institutionalized population is defined as the population belonging to households whose heads are not inpatients of a hospital, an inmate of a social institution, a person in a Self-Defence Forces camp, an inmate of a reformatory institution, or other social institution. The analysis sample includes only household members who "mainly worked". Column (6) tabulates the analysis sample size used to delineate the CZs.

respondent's work location since 1980, and this is planned to continue, whereas the US Population Census ceased to ask that question in 2010. This has required American researchers to find other sources of workplace data, and the American Community Survey (ACS), a sample survey and not a census, has been adopted for this purpose. However, in the Fowler et al. (2016) reassessment of US CZs in 2010 using ACS data, they find significant variations in the results depending on the application method, so the consistent data source available for CZ research in Japan is a non-negligible benefit. Another benefit of using a census rather than a sample is to have information on commuting flows without margins of error (MOE). In their discussion of the role of MOEs in determining CZs in the US, Foote et al. (2021) state that MOEs distort the dissimilarity measures between counties in a non-linear fashion, making it difficult to apply the law of large numbers to remove the errors. As the Japanese Population Census does not have sampling errors, our clustering is free from such errors.

The strength of CZs as compared to UEAs is that the derivation process is data-driven and largely free from any prior knowledge of geographic divisions except for the choice of the cutoff value. To delineate UEAs, we need to specify the center of the employment area around which the employment area is defined. This first step entails the choice of the threshold of the population density that generally requires the expert knowledge so that the resulting delineation makes sense. In addition, the recent substantial improvement of computing power makes the implementation of Tolbert and Sizer (1996) even more free from the expert knowledge than the original process proposed by Tolbert and Killian (1987) and Tolbert and Sizer (1996). The original implementation uses expert knowledge to separate counties into two or more regions in the US to reduce the computational burden. Fowler et al. (2016) however point to a potential bias in the resulting delineation. We simply input the full commuting matrix into the algorithm to avoid any possibility of this potential bias.

As for general weaknesses of the method, Fowler et al. (2016) point out the following three, all of which also apply to our implementation. First and foremost, it tends to privilege the connection between large and small municipalities, since the denominator of the (dis)similarity measure is determined by the size of the small municipality. In other words, the method undervalues the connection between two large municipalities that would have large denominators in the similarity measure (so the measure overstates the dissimilarity). This may imply that the agglomerated large metropolitan areas such as those centered around Tokyo are likely to be separated from each other. Second, the average linkages method of calculating the distance between clusters undervalues the connection between one large core and many peripheries since, for each periphery municipality, the distance to the rest of the cluster also reflects the commuting distance to other peripheries and not only to the core. Finally, we do not have external measures to evaluate the goodness of fit. This last point is discussed more fully in Section 5.

To conclude our methodological discussion, we point out that Fowler et al. (2016) also compare the performance of CZs with other labor market concepts according to how connected and contained is each delineated region. According to their analysis, the CZ is generally more connected but less contained than other available measures in the US that are mutually exclusive and exhaustive. Our future research is to study these and other measures to compare the performance of various labor market concepts in Japan.

4 Results

The CZs constructed according to the algorithms described in Section 2 are summarized in Table 2. One thing we should note is that the number of municipalities in Japan has decreased over time, particularly during the 2000s, due to mergers (Weese, 2015). The results in column (1) show the number of CZs derived based on the number of municipalities (column (2)) as defined at each Census year. For comparison, the results in columns (3) and (4) are based on a harmonized municipality code as of 2015.

For 2015, 265 CZs were derived from 1,736 municipalities, so the average number of municipalities per CZ is 6.55. In the US, Tolbert and Sizer (1996) obtain 741 CZs from 3,141 counties, which results in an average count of 4.24 counties per CZ. The larger number of average municipalities in Japan is not surprising given the smaller size of municipalities in Japan as compared to US counties and the greater mobility across smaller local units.

We now examine the properties of the derived CZs. in the interest of space, we discuss only 2015 here, but the discussion applies to other years as well. Figure 1 is a colored map that shows the delineation of our output CZs based on the input of the harmonized municipality codes in 2015. Panel (a) shows the nationwide map while Panel (b) shows a close-up of Kanto district around Tokyo.[6] The figures show that CZs are delineated as geographically contiguous units even though our algorithm outlined in 3 does not impose the restriction that the municipalities in a CZ must be contiguous. In particular, if many workers in a municipality commute to another municipality

[6]Appendix B.1 shows the close-up of other metropolitan regions, Osaka and Nagoya.

Table 2: Number of CZs in Each Population Census

Municipality code	Original		Harmonized	
Year	CZs	Municipalities	CZs	Municipalities
	(1)	(2)	(3)	(4)
1980	684	3,278	405	1,741
1985	609	3,276	367	1,741
1990	542	3,268	334	1,741
1995	496	3,255	315	1,741
2000	450	3,251	292	1,740
2005	322	2,239	277	1,741
2010	267	1,750	267	1,741
2015	265	1,736	265	1,736
2020	258	1,740	258	1,740

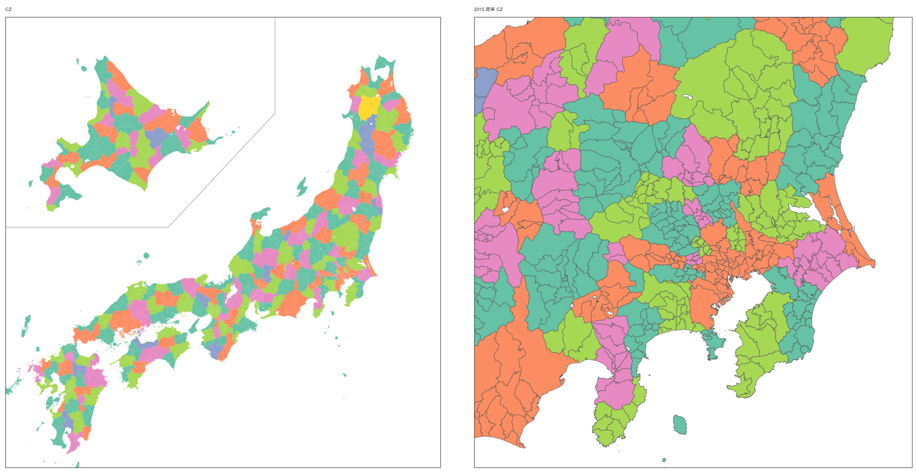
Note: This table shows the results of the Hierarchical Agglomerative Clustering method described in section 3.2. The "Original code" columns are the results of delineating the commuting zones using the definition of the municipality as of each survey year. The "Harmonized code" columns are based on the definition of municipalities as of 2015. In 2000, one municipality was excluded from the analysis due to the impact of a volcanic eruption. Similarly, in 2015, five municipalities were excluded from the analysis due to the effects of nuclear damage. A City with multiple administrative districts, called a government-designated city (*Seirei Shitei Toshi*), is treated as a single city.

that is not contiguous then geographically separated municipalities could form a CZ that entails enclaves. Merger of non-contiguous municipalities is relevant when the residential population of central business district (CBD) is smaller than the commuter to the CBD from a municipality, because the distance metric between the CBD and the municipality is zero. This case indeed happens for Chiyodaku of Tokyo-to, a CBD of Tokyo, resided by a small population. Given this, a notable feature from this figure is that, there are no enclaves in any of the derived CZs. This is particularly striking given Japan's intensive use of trains for commuting and the fact that housing development tends to cluster most densely around stations. We interpret this in the following three ways. First, the commuting pattern may generally be characterized by geographic proximity, so that even while good station-based commuting connections may well play an important role, it does not override the effect of geographic proximity. Second, our clustering method may classify two station-centered economic zones as separate CZs. Third, the threshold is high enough so that enclaves generated in the middle of the process are merged into a geographically contiguous CZ. In any case, we view it as an attractive feature that we do not have any enclaves in our results, as the analysis of local labor markets containing enclaves is not trivial. Further, this feature is shared by other major local labor market definitions such as the US CZs (Tolbert and Sizer, 1996) or Japan's

UEA (Kanemoto and Tokuoka, 2002).

Figure 1: 2015 CZ Delineation

(a) Nationwide (b) Kanto region

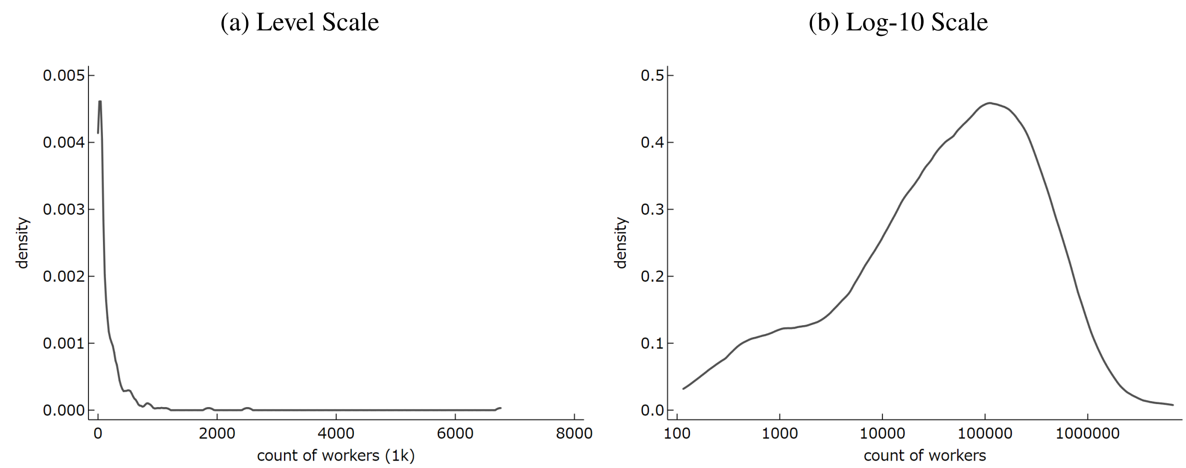


Note: This map is based on the authors' calculations using the 2015 Population Census and hierarchical agglomerative clustering with average linkages. Harmonized municipality codes were used as input for the algorithm. The ESRI shapefile used to create this map was developed by Kirimura et al. (2011).

To further characterize the properties of our CZs, Figure 2 shows the distribution of workers in each CZ as estimated by the kernel density method. As can be seen from the estimated density, there are many small-scale CZs and a few large-scale CZs, and this skewed distribution is especially apparent in the log-scaled distribution shown in the panel (B). Reflecting the right skewed distribution, the sample mean is 175,683 while the median is 53,607. The sample standard deviation is 483,121.

To assess the temporal consistency of our CZ definitions, we conducted a graphical analysis of CZ stability across census years from 1980 to 2015. This analysis reveals boundary consistency throughout the study period, with many areas maintaining stable CZ affiliations over time. In many cases, CZ forms change over time only when multiple CZs are consolidated. These shifts generally reflect actual changes in commuting patterns rather than methodological inconsistencies. The temporal stability of our CZ definitions supports their reliability for longitudinal research spanning multiple decades.

Figure 2: Distribution of the Number of Workers



Note: Figure of the kernel density estimates based on authors' calculation based on the 2015 Population Census. The density is estimated with the Epanechinikov kernel function and optimal bandwidth.

4.1 Choosing the Tree Height

Table 3 presents our analysis of how different dissimilarity cutoff values affect the explanatory power of CZ definitions for log hourly wages. Each column progressively adds fixed effects for clusters defined at increasingly granular levels of tree height. As we move from larger to smaller clusters (higher to lower tree heights), we observe a gradual increase in R-squared values, indicating that more granular CZ definitions explain additional variation in wages. However, this improvement comes with diminishing returns, particularly below the 0.98 threshold. The F-statistics demonstrate that each level adds statistically significant explanatory power, but the magnitude of improvement decreases as clusters become smaller. This analysis supports our choice of 0.98 as the baseline cutoff, as it balances explanatory power with practical usability. Further discussion of the threshold choice is made in Section 5.

4.2 Performance Assessment

As our goal is to define the geographic units that best represent the local labor market, it is important that we evaluate the goodness of fit of our derived CZs. We evaluated the performance of our derived CZs as compared to other potential geographic units in two ways: 1) the fraction of workers who commute within CZs and 2) an analysis of variance (ANOVA) with respect to labor market outcomes.

Table 3: Analysis of Variance across Tree Heights

VARIABLES	(1) ln_hwage	(2) ln_hwage	(3) ln_hwage	(4) ln_hwage	(5) ln_hwage	(6) ln_hwage
Added Dummies F-statistic	41.322	7.432	9.299	5.722	4.877	4.513
Observations	412,742	412,742	412,742	412,742	412,742	412,742
R-squared	0.150	0.151	0.153	0.153	0.154	0.155
Cluster h=0.99 Dummies	✓	✓	✓	✓	✓	✓
Cluster h=0.98 Dummies	-	✓	✓	✓	✓	✓
Cluster h=0.97 Dummies	-	-	✓	✓	✓	✓
Cluster h=0.96 Dummies	-	-	-	✓	✓	✓
Cluster h=0.95 Dummies	-	-	-	-	✓	✓
Cluster h=0.94 Dummies	-	-	-	-	-	✓
N of Clusters	177	247	295	341	414	460
Added Dummies p-value	0.000	0.000	0.000	0.000	0.000	0.000

Note: The result of ANOVA analysis with respect to tree height h. Each column iteratively adds a finer layer of clusters defined by each level of the tree height indicated by checkmarks. The dependent variable is log hourly wage. The demographic controls are five-year age bins, sex dummy, and four-bin educational attainment. All regressions are weighted by the sampling weight of ESS.

First, we examined the proportion of workers who commute within a geographic unit. Figure 4 shows the box diagrams of the distributions when the unit is a municipality, CZ, or prefecture. Panels (a) and (b) show the results unweighted and weighted by population, respectively. We can see that when the municipality is the unit of analysis, the proportion of commuters is below 0.7, meaning that more than 30% of workers work outside of the municipality. By contrast, when the CZ or prefecture is the unit of analysis, the proportion increases up to to 0.95, meaning that fewer than 5% of workers commute to a workplace located outside of the unit. A notable difference between the population unweighted and weighted results for both prefectures and CZs is the difference in the 25th percentile, with the unweighted results providing a higher proportion of commuters within the unit. Comparing CZs and prefectures as the unit of analysis, the performance is similar, and while enlarging the geographic area to the prefecture slightly increases the proportion of within-unit commuters, it comes at the cost of a substantial decrease in the number of geographic units. In sum, the analysis demonstrates that the proposed CZs capture actual commuting patterns well while keeping the number of observations reasonably high.

We have extended our performance evaluation across all census years from 1980 to 2020, replicating the catchment rate and ANOVA analyses for each period. This comprehensive assessment reveals remarkably consistent CZ performance over time, with only minor variations attributable to structural changes in Japan's labor markets. These results confirm that our CZ definitions maintain their utility for analyzing local labor markets across different time periods. See Appendix Figure B.3 for more details.

As an additional performance assessment, we examined how much the local labor market is integrated within various geographic units by testing the law of one price within a single market. Specifically, if workers in a local labor market can find jobs without friction, then every worker will face the same offered wage, and identical workers, including their reservation wages, are expected to share the same market outcomes such as employment and wages. To check how integrated the labor market is, we examined how much of the variation in labor market outcomes is shared across individuals within a geographic unit, conditional on demographics. To study how each CZ captures the variation in these data, we conducted an analysis of variance drawing on the 2017 Employment Status Survey (ESS), administered by the Ministry of Internal Affairs and Communications, and extracted labor market variables (working or not and annual income), demographic variables (age, sex, education), and the municipality of residence. Using this, we matched our CZ to the municipality.[7] Demographic variables are five-year age bins, sex dummy variable, and four-bin educational attainment (Less than high school diploma, High school diploma, Technical/Vocational school diploma, four-year university diploma). We then studied how the variation in the labor market variables could be explained by different layers of geographic units from coarse to fine; namely, prefectures, CZs, and municipalities, after conditioning on the demographic variables. We implemented this by adding the fixed effects (FEs) of each layer of geographic units to the linear regressions. The regression is weighted by the sampling weights of the ESS. Formally, we ran the following regression:

$$y_i = \alpha_{l(i)} + X_i\beta + \varepsilon_i, \tag{1}$$

where i is individual and $l(i)$ is the location of i. The vector X_i includes the demographic variables: five-year age bins, sex dummy, and four-bin educational attainment. We are interested in the role of different layers of geographic fixed effects $\alpha_{l(i)}$, measured by R-squared and the F statistics of added fixed effects.

[7] The choice of 2017 ESS is simply based on its proximity to the 2015 Population Census. ESS is surveyed every fifth year, ending with digits 2 and 7. Thus the 2017 ESS is closest to capturing the labor market patterns as of 2015.

Table 4 and 5 show the result of the analysis of variance for the variable of employment and log earnings, respectively. Column 1 of Table 4 shows that the demographic variables explain 0.388 of the total variation. Adding 47 prefecture FEs increases the fraction up to 0.389, and further addition of 227 CZ FEs increases the fraction up to 0.390. Finally 1372 municipality fixed effects renders the fraction 0.393. Reflecting the large sample size, the additional fixed effects are all statistically significant in F-tests. Among the changes, the relatively small value of F statistics when we control CZ FEs relative to municipality FEs is notable reflecting the large decrease in the degree of freedom due to the large number of FEs added to the regression. We find similar tendency in the F-test results for annual earnings reported in Table 5. Overall, the CZs capture significant amount of variations of the employment and earnings within a prefecture. Municipalities captures the additional variation within a CZ but additional gain is limited relative to the decrease in the degree of freedom.

Table 4: Employment

VARIABLES	(1) emp	(2) emp	(3) emp	(4) emp	(5) emp	(6) emp
R-squared	0.393	0.394	0.396	0.396	0.396	0.399
Added FE F statistics	-	38.34	6.65	6.65	6.65	3.06
Demographic Controls	✓	✓	✓	✓	✓	✓
Prefecture FE		✓	✓	✓	✓	
CZ (h=0.99) FE			✓			
CZ (h=0.98) FE				✓		
CZ (h=0.97) FE					✓	
Municipality FE						✓

Note: The table presents the results of the ANOVA analysis in equation (1). The dependent variable is the dummy variable indicating working (including part time). When including municipality FEs, prefecture and CZ FEs are not identified since municipality delineations are contained in those of prefectures and CZs. The reported number of added FEs in each column is the marginal addition of FEs in each column relative to the one left, netting out the non-identified FEs in the other FE categories. The demographic controls are 5-year bin age dummies, sex dummies, and 6-level education attainment dummies (Less than Middle School, High School, Business College, Junior/Technical College, 4-year University, Graduate School). The sample size is 787,466 and all regressions are weighted by the sampling weight of ESS. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Log Annual Earnings

VARIABLES	(1) l_inc	(2) l_inc	(3) l_inc	(4) l_inc	(5) l_inc	(6) l_inc
R-squared	0.326	0.332	0.335	0.335	0.335	0.340
Added FE F statistics	-	52.67	4.68	4.68	4.68	2.14
Demographic Controls	✓	✓	✓	✓	✓	✓
Prefecture FE		✓	✓	✓	✓	
CZ (h=0.99) FE			✓			
CZ (h=0.98) FE				✓		
CZ (h=0.97) FE					✓	
Municipality FE						✓

Note: The table presents the results of the ANOVA analysis from equation (1). The dependent variable is the natural logarithm of the median value of earnings classification in the ESS survey. For the largest annual earnings category, 15 million JPY or above, we assigned 15 million JPY. When including municipality FEs, prefecture and CZ FEs are not identified since municipality delineations are contained in those of prefectures and CZs. The reported number of added FEs in each column is the marginal addition of FEs in each column relative to the one left, netting out the non-identified FEs in the other FE categories. The demographic controls are 5-year bin age dummies, sex dummies, and 6-level education attainment dummies (Less than Middle School, High School, Business College, Junior/Technical College, 4-year University, Graduate School). The sample size is 455,304 and all regressions are weighted by the sampling weight of ESS. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.3 Comparison with UEAs

We also compare the performance of our 2015 CZs with 2015 Urban Employment Areas (UEAs) in Figure 5. While UEAs focus on urban hierarchies using population thresholds, CZs capture actual commuting patterns without arbitrary size restrictions. This results in more comprehensive coverage of rural areas, better detection of emerging employment clusters, and greater flexibility in capturing complex commuting relationships. Our comparative analysis helps users make informed decisions about which regional definition best suits their research needs.

We also compare the CZs and UEAs in 1980. As expected, the CZs and UEAs are coincident, as in 2015. However, the number of regions included in the UEA is smaller in 1980, making the choice of UEA years more arbitrary for researchers. The details are shown in Appendix Figure B.2.

4.4 Latest Development

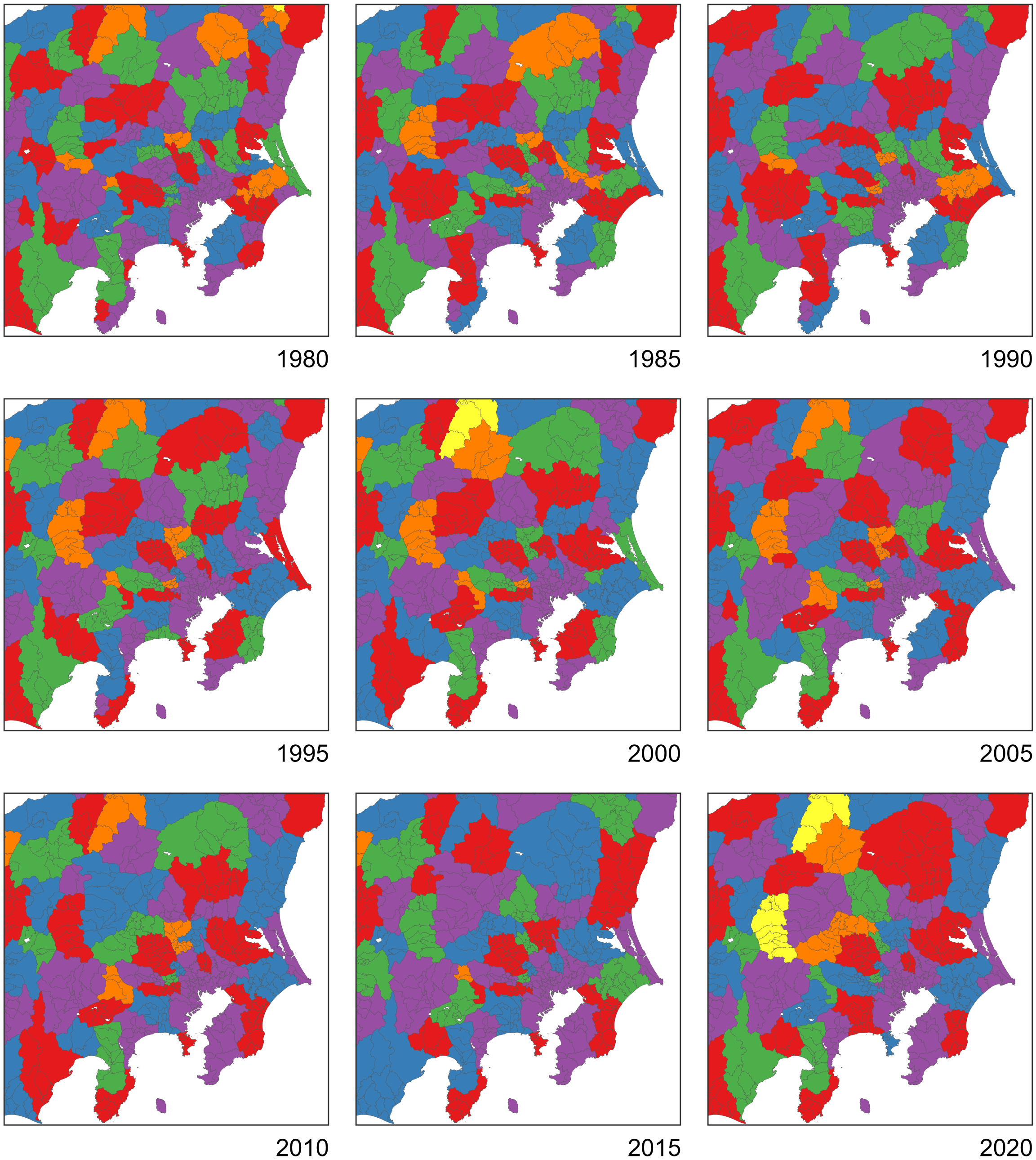
We acknowledge that recent developments, particularly the COVID-19 pandemic, may significantly alter commuting patterns through the rise of remote work. While our current CZ definitions are largely based on pre-pandemic data, we propose an adaptive methodological approach that will allow for recalibration as new data becomes available. Future updates to these CZ definitions should incorporate changes in commuting distances, frequencies, and the prevalence of remote work to ensure ongoing relevance in a changing labor market landscape. We also show CZ delineation from 2020 to discuss the potential effect of the pandemic.[8]

5 Discussion of Dissimilarity Cutoff

In this section, we further discuss key considerations regarding our choice of the dissimilarity cutoff, or tree height, which is a critical hyperparameter of the HAC method. Figure 6 displays the complete dendrogram of our clustering method. The "Height" on the y-axis represents the tree height at which each node (or group of nodes) merges into a larger group, and it is evident that the tree branches become dense near the cutoff $D \approx 1$. The red rectangle highlights the metropolitan Tokyo CZ. The blue dashed line indicates our selected cutoff value of 0.98.

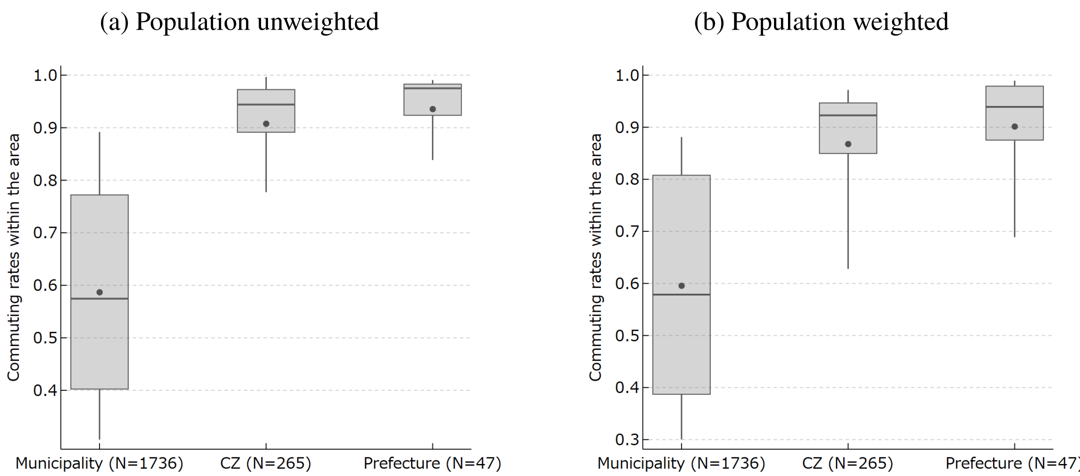
[8] A particularly notable example demonstrating the robustness of our methodology is the evolution of CZs in Fukushima Prefecture from 2010 to 2020. Our methodology allows us to include areas affected by residential restrictions due to the nuclear accident following the 2011 Great East Japan Earthquake, yielding meaningful results. Specifically, we observed that regions forming a single CZ in 2010 were divided into different CZs in 2015 (with commuting flows still observed, such as for decontamination work), and then reintegrated into a single CZ in 2020 following the partial lifting of residential restrictions. This case demonstrates that even when faced with abrupt changes due to exogenous shocks, our CZ delineation methodology captures substantive relationships between regions and enables temporally consistent analysis.

Figure 3: CZ evolution over time (Kanto Region)



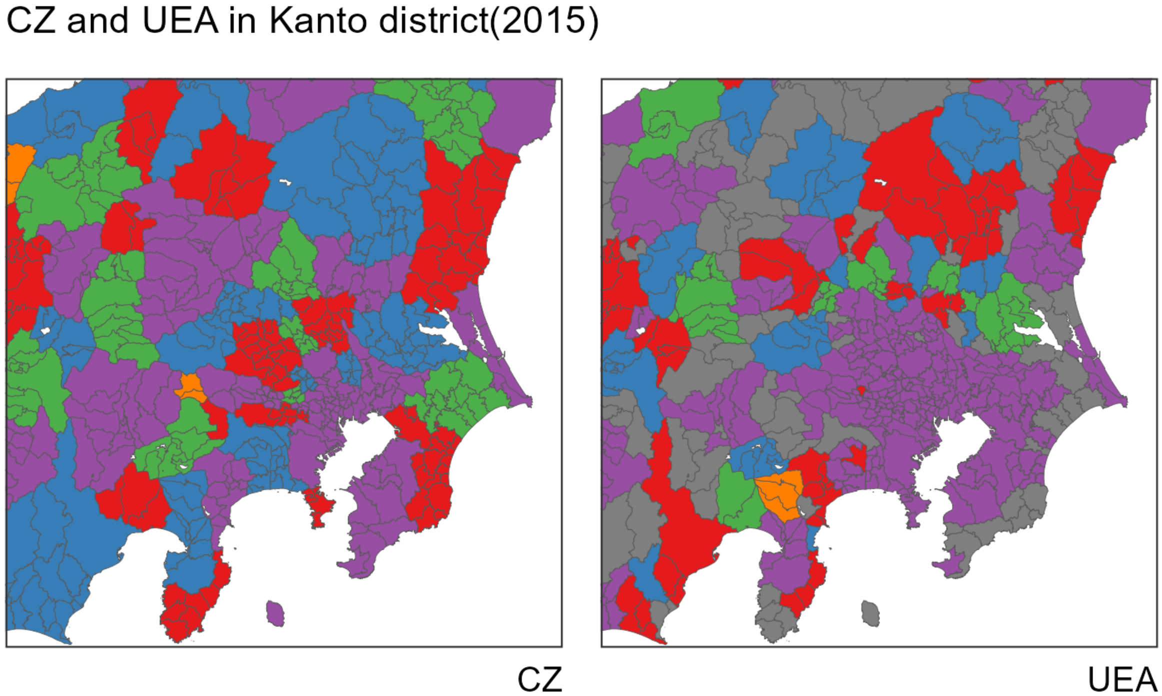
Note: This figure illustrates the evolution of Commuting Zone boundaries in the Kanto region from 1980 to 2020, based on the Population Census data. The maps use harmonized municipality codes as of 2015. Different colors represent distinct CZs. The CZ that contains Chiyoda, Tokyo maintains its color throughout the period, while other CZs may have different colors across years. The ESRI shape file for creating the map was generated by Kirimura et al. (2011).

Figure 4: Proportion of Workers Who Commute within a Geographic Unit



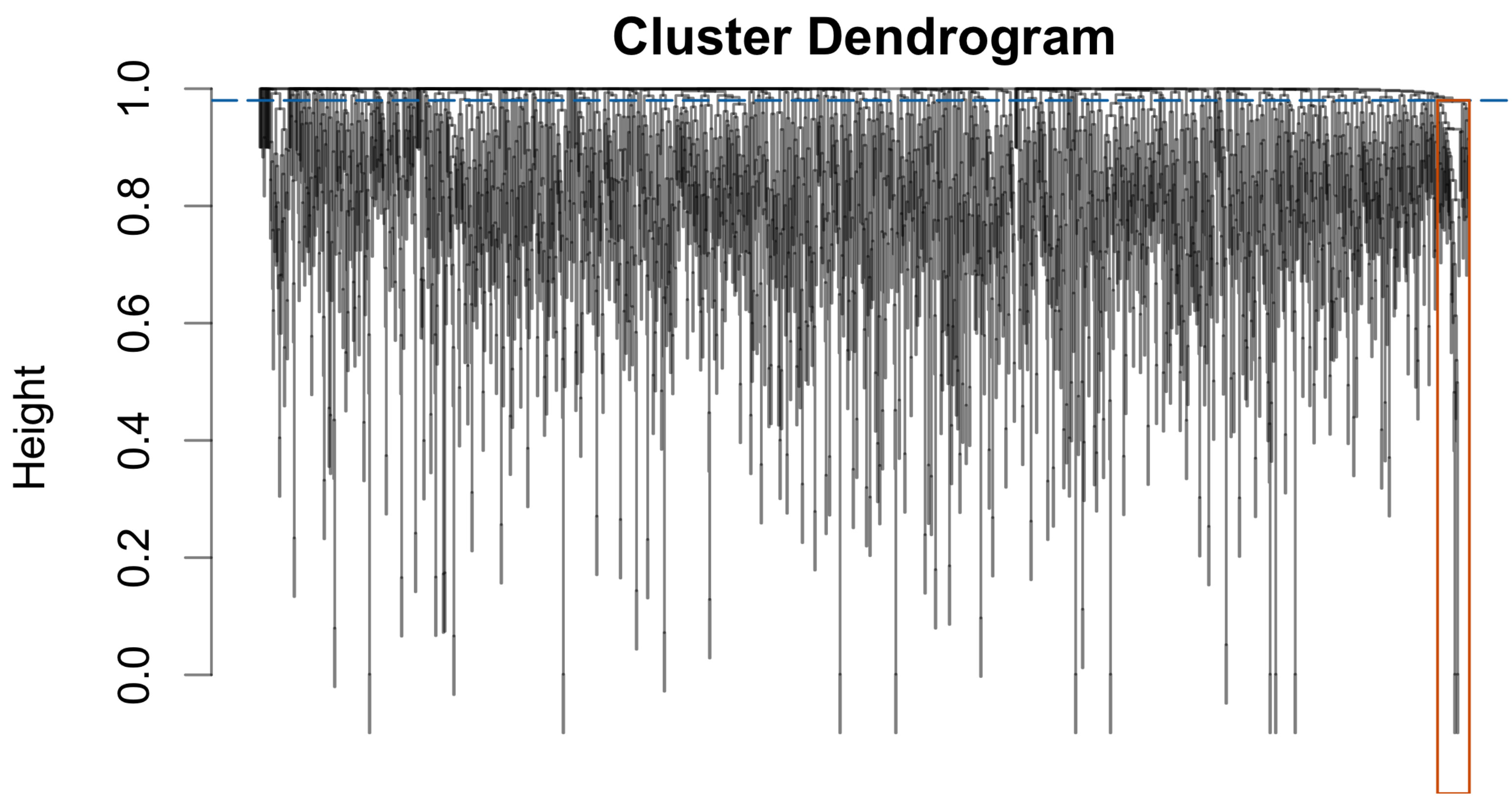
Note: This figure shows the distribution of the share of residents who commute within their own area, measured at the levels of prefectures, commuting zones, and municipalities in 2015. The dot indicates the mean, the line in the middle of the box indicates the median, and the upper and lower ends of the box indicate the 75th and 25th percentiles. The line above and below the box show the 90th and 10th percentiles. Panel (a) shows the unweighted results and panel (b) shows the weighted results.

Figure 5: CZs and UEAs in Kanto region in 2015



Note: The left panel shows the UEAs and the right panel shows the CZs. Gray color shows the area that is not included in UEAs.

Figure 6: Overall Dendrogram



Note: The figure shows the overall dendrogram with the HAC method applied to Japan's 2015 Population Census. The red rectangle represents the metropolitan Tokyo CZ, which is detailed in Figure 7. The horizontal blue dashed line indicates our cutoff value, 0.98.

Figure 7 shows a close-up of the dendrogram for the metropolitan Tokyo CZ. One can see from the order of the node labels that indicate municipality names, the relatively eastern wards in Tokyo and beyond are on the top, while western Tokyo wards and beyond are on the bottom. For example, in the middle, "Tokyo-to, Koutouku" is merged with "Tokyo-to, Chiyodaku" at tree height 0.5. This happens because a large proportion of residents of Koutouku, a residential area, commute to Chiyodaku, which is the Central Business District with relatively few residents. Then, to this Koutouku-Chiyodaku cluster, the Minatoku-Chuuouku cluster is merged, as large numbers of its residents also commute to Chiyodaku.

As these figures illustrate, the critical question is at which height of the dendrogram we should stop merging municipalities. In Figure 6, the blue dashed line indicates our cutoff value of 0.98, and we find that the tree branches become dense near $D \approx 1$, which implies that the number of CZs may be sensitive to changes in the dissimilarity cutoff h close to 1.[9] For the remainder of this section, we aim to quantify this sensitivity.

To see how sensitive our definition of CZs is to the choice of dissimilarity cutoff h, we show in Figure 8 the plots of cluster counts as a function of h. The left and right panels show plots for $h \in [0.5, 1]$ and $h \in [0.9, 1]$, respectively. In each figure, counts of relevant geographic units are displayed with dotted horizontal lines, and our choice of tree height $h = 0.98$ is represented by a dashed vertical line. For example, one can see that for our algorithm to obtain 741 CZs in Japan, the same number as was found in the US, our tree height would have needed to be $h \approx 0.90$, while to reach the 94 UEAs found in Japan, a $h \approx 0.99$ would have been required. As discussed above, there are reasons why the number of CZs is larger in the US than in Japan, while there are fewer UEAs than CZs in Japan. As these can be considered the upper and lower bounds, it is reasonable that our choice of $h = 0.98$ lies between these values.

We have demonstrated that the dissimilarity cutoff point (tree height) is inversely related to the number of clustering units, so setting a high cutoff reduces the number of CZs when we implement the procedure. On the other hand, the benefit of setting a high cutoff is an increase in the proportion of workers who actually commute within the CZ. Panel (a) of Figure 9 illustrates the relationship between tree height and proportion of workers commuting within clusters. It clearly shows that an increase in tree height, implying an enlargement of the commuting zone, increases the probability of properly catching commuters. Our choice of a tree height of 0.98 captures 87% of the actual commuting pattern. This means that the remaining 13% commute to a location outside of our

[9] In Figure 7, the blue line indicates the cutoff value of 0.98. One can confirm that each pair of nodes are combined at a lower value than the cutoff since all municipalities in the figure are in one CZ.

defined commuting zone, so this is considered catchment error. This catchment error can be reduced by increasing the tree height, as demonstrated in Panel (a), but the increase in tree height would also increase the risk of merging heterogeneous labor markets that do not belong to the same local labor market. This is the tradeoff in choosing the cutoff. Next, as a way to capture the heterogeneity of labor market outcomes within a cluster, we calculated the $1 - R^2$ (Error variance / Total variance) of the regression of employment status on geographic fixed effects corresponding to the tree height, five-year age bins, sex dummy variable, and four-bin educational attainment as we did in the Anova analysis in equation (1). This $1 - R^2$ captures the heterogeneity in the probability of employment among people who share the same demographic variables within a cluster. Panel (b) of Figure 9 show the results and, as expected, as tree height increases, so does the heterogeneity of employment outcomes within a cluster. Thus, we have shown that enlarging the geographic unit by increasing the tree height increases both the catchment rate and employment heterogeneity, creating a trade off between the two factors when choosing the cutoff.

Finally, we discuss the applicability of cross validation method to best determine the tree height (among others, these include Ben-Hur et al., 2001; Tibshirani and Walther, 2005; Fang and Wang, 2012, Fu and Perry, 2019). These pre-existing methods handle the situation that the eventual number of clusters is substantially smaller than our case. For example, Fang and Wang (2012) discusses a method of using cross validations to pin down the cluster count to a range of 2 to 10. In contrast, our goal is to generate more clusters than the scope of these papers. When we applied the cross validation method to our setting, allowing for a three-digit cluster count to accommodate the 741 CZs found in the US by Tolbert and Sizer (1996), the process did not converge in a realistic time. Thus the optimal choice of tree height using the cross validation method remains a subject for future study.

In sum, our analysis articulates the trade-offs between the number of clusters, catchment rate, and heterogeneity within a cluster. Setting a lower value for the dissimilarity cutoff increases the number of CZs and homogeneity within the CZ, but decreases the catchment rate. Determining the optimal cutoff height requires a reasonable objective function that is maximized with the trade-off as a constraint. To the best of our knowledge, the choice of the objective function is not discussed in the CZ construction literature and is left for future research. Our choice of 0.98 as the cutoff based on Tolbert and Killian (1987) and Tolbert and Sizer (1996) is at least justified for its international comparability to the US.

6 Final Remarks

In this paper, we delineated mutually exclusive and exhaustive Commuting Zones (CZs) in Japan using the Hierarchical Agglomeration Clustering (HAC) method and Japan's Population Census data from 1980 to 2020, obtaining from 265 to 684 unique CZs depending on the Census year. The performance of the resulting CZ delineations was then assessed in comparison to other potential geographic units commonly used for economic analysis.

Our approach to defining commuting zones can be thought of as the application of a modern spatial theory while remaining empirically flexible (Allen and Arkolakis, 2025). Rather than relying on restrictive assumptions of monocentric city models, we allow labor market areas to emerge organically from observed commuting patterns. This data-driven approach enables us to capture the complex, multi-directional nature of commuting flows that characterize modern Japanese cities. The thresholds emerge from data analysis guided by spatial economics principles and research linking commuting patterns to economic integration indicators like wage convergence and employment distribution.

We find that, rather than a "magic bullet" to be used in all situations, the proposed CZ is a geographic unit that can be useful for many practical purposes. For example, the 265 CZs derived in 2015 increases the number of observations as compared to the 47 prefectures while also better capturing the actual commuting patterns of Japan's 1,736 municipalities. This fact can help to better understand any causal relationship that goes beyond municipal boundaries, such as a shock to an establishment located in one municipality which affects a commuting worker living in another municipality. In this case, using the municipality as the geographic unit of analysis fails to capture the causal relationship. An additional benefit of CZs is that the statistics calculated on CZ suffer less from small sample problems. In the end, as always, the choice of geographic unit of analysis should be based on the pros and cons of each potential geographic unit depending on the purpose of the analysis at hand, and so with this paper we aim to add a useful device to the empirical researcher's toolbox.

Finally, we present a caution to readers that CZs are time varying depending on changes in commuting patterns, which we found in the various waves of the Population Census. What this means is that, first, updates of CZ delineation is required continuously. This work is straightforward with the program code used in the current analysis and the municipality level matrix with the number of commuters between origins and destinations. The codes for CZ delineations for the Population Census from 1980-2020 are available on GitHub (`https://github.com/daisukeadachi/`

`commuting_zone_japan`). Second, when data from multiple years are pooled, researchers need to choose the CZ delineation year carefully, depending on the purpose of the analysis. Although it is difficult to exhaust the mapping from the purpose to the actual choice of CZ year, we have provided general guidance for the choice strategy in section C of the Appendix.

Statements and Declarations

All authors declare that they have no financial or non-financial competing interests related to this manuscript.

References

Acemoglu, Daron and Pascual Restrepo, "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

Allen, Treb and Costas Arkolakis, "Quantitative Regional Economics," January 2025.

Alonso, William et al., *Location and land use. Toward a general theory of land rent.*, Cambridge, Mass.: Harvard Univ. Pr., 1964.

Altonji, Joseph G. and David Card, "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in "Immigration, Trade, and the Labor Market" NBER Chapters, National Bureau of Economic Research, Inc, October 1991, pp. 201–234.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–2168.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute for Employment Research, 1991.

Ben-Hur, Asa, Andre Elisseeff, and Isabelle Guyon, "A stability based method for discovering structure in clustered data," in "Pacific Symposium on Biocomputing," World Scientific, 2001, pp. 6–17.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, " Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 09 2014, *129* (4), 1553–1623.

Cörvers, Frank, Maud Hensen, and Dion Bongaerts, "Delimitation and coherence of functional and administrative regions," *Regional studies*, 2009, *43* (1), 19–31.

Dorn, David, "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD dissertation, Verlag nicht ermittelbar 2009.

Fang, Yixin and Junhui Wang, "Selection of the number of clusters via the bootstrap method," *Computational Statistics & Data Analysis*, 2012, *56* (3), 468–477.

Foote, Andrew, Mark J Kutzbach, and Lars Vilhuber, "Recalculating...: How Uncertainty in Local Labor Market Definitions Affects Empirical Findings," *Applied Economics*, 2021, *53* (14), 1598–1612.

Fowler, Christopher S, Danielle C Rhubart, and Leif Jensen, "Reassessing and Revising Commuting Zones for 2010: History, Assessment, and Updates for US 'Labor-Sheds' 1990–2010," *Population Research and Policy Review*, 2016, *35* (2), 263–286.

Fu, Wei and Patrick O Perry, "Estimating the number of clusters using cross-validation," *Journal of Computational and Graphical Statistics*, 2019, pp. 1–12.

Fujita, Masahisa, Tomoya Mori, J. Vernon Henderson, and Yoshitsugu Kanemoto, "Chapter 65 - Spatial Distribution of Economic Activities in Japan and China," in J. Vernon Henderson and Jacques-François Thisse, eds., *Cities and Geography*, Vol. 4 of *Handbook of Regional and Urban Economics*, Elsevier, 2004, pp. 2911 – 2977.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Kanemoto, Yoshitsugu and Kazuyuki Tokuoka, "Metropolitan Area Definitions in Japan (In Japanese)," *Journal of Applied Regional Science*, 2002, *7*, 1–15.

Kirimura, T, T Nakaya, and K Yano, "Building a Spatio-temporal GIS Database about Boundaries of Municipalities: Municipality Map Maker for Web," *Theory Appl GIS*, 2011, *19*, 83–92.

Kropp, Per and Barbara Schwengler, "Stability of functional labour market regions," Technical Report, IAB-Discussion Paper 2017.

Mills, Edwin S, "An aggregative model of resource allocation in a metropolitan area," *The American Economic Review*, 1967, *57* (2), 197–210.

Muth, Richard F, *Cities and Housing; The Spatial Pattern of Urban Residential Land Use*, University of Chicago Press, 1969.

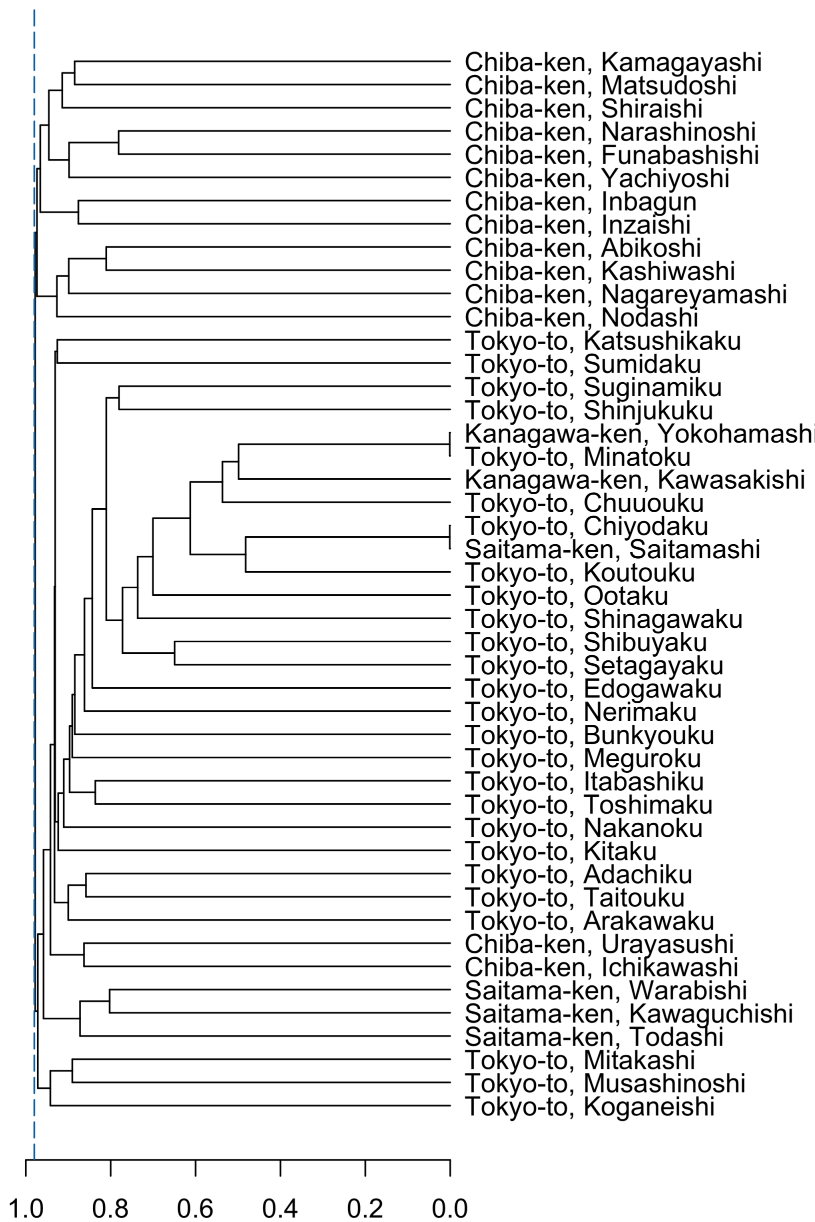
Tibshirani, Robert and Guenther Walther, "Cluster validation by prediction strength," *Journal of Computational and Graphical Statistics*, 2005, *14* (3), 511–528.

Tolbert, Charles and Molly Killian, "Labor Market Areas for the United States. Washington, DC: Department of Agriculture," *Economic Research Service, Agricultural and Rural Economics Division*, 1987.

Tolbert, Charles M. and Molly Sizer, "U.S. Commuting Zones and Labor Market Areas: A 1990 Update," Technical Report, ERS Staff Paper 1996.

Weese, Eric, "Political mergers as coalition formation: An analysis of the Heisei municipal amalgamations," *Quantitative Economics*, 2015, *6* (2), 257–307.

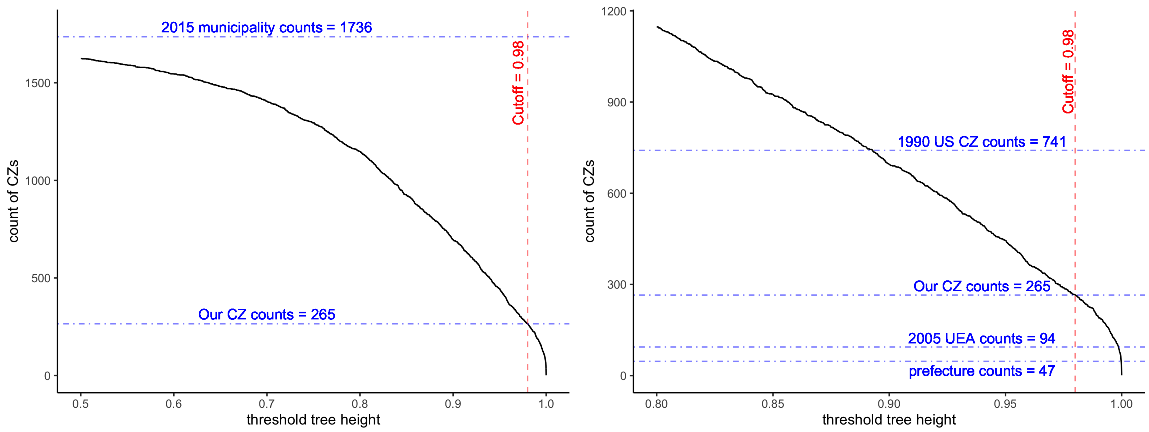
Figure 7: A Focus on Metropolitan Tokyo CZ



Note: This figure shows a close-up of Metropolitan Tokyo CZ from the overall dendrogram in Figure 6. The overall dendrogram was generated using the HAC method and 2015 Census data as discussed in the main text. This figure is rotated 90-degrees counterclockwise to make the labels more readable. The vertical blue dashed line indicates our cutoff value, 0.98. This figure includes the cases that municipalities are merged at height zero because of the small residential population of one municipality. For example, the nighttime population of Tokyo-to, Chiyodaku was 17,369 and the numbers of commuters from Saitama-ken, Saitamashi was 24,051, thus the distance metric is zero. The same thing happens for Kanagawa-ken, Yokohamashi and Tokyo-to, Minato-ku.

Figure 8: Dissimilarity Cutoff and Number of CZs

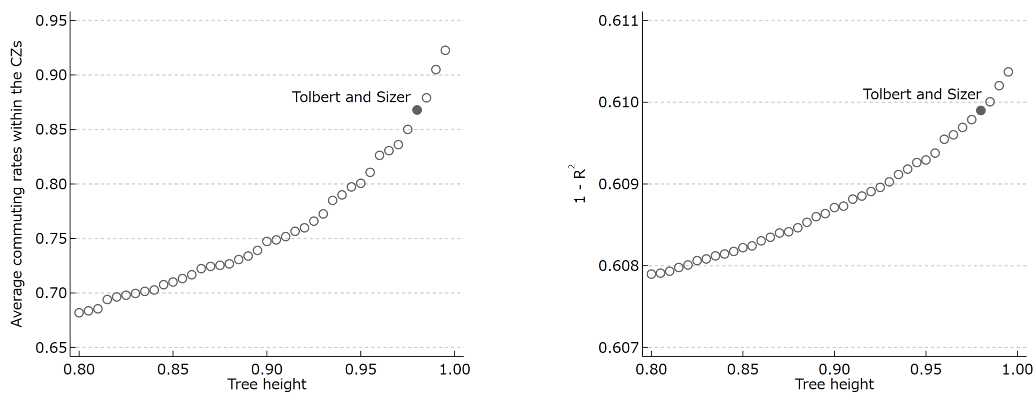
(a) Cutoff values between 0.5 and 1 (b) Cutoff values between 0.8 and 1



Note: Figure was constructed by authors based on the 2015 Japan Population Census. Cutoff values with a bin of 0.0001 were clustered according to the method described in the main text and the count of the resulting CZs was plotted. The red vertical line highlights our preferred choice, 0.98, based onTolbert and Killian (1987). The blue horizontal lines provide reference values of relevant geographic units.

Figure 9: Cutoff Point, Within-Cluster Commuting Rate, and Heterogeneity in the CZ

(a) Within-Cluster Commuting (b) 1-R^2 of Employment Equation



Note: Figure was constructed by authors based on the 2015 Japan Population Census. Cutoff values with a bin of 0.01 were clustered according to the method described in the main text and the count of the resulting CZs was plotted. The red vertical line highlights our preferred choice, 0.98, based onTolbert and Killian (1987). The blue horizontal lines provide reference values of relevant geographic units.

Appendix

A CZ Correlates with Other Socioeconomic Variables

A.1 Density and Industry Share

In this section, we show the results of our analysis based on the properties of CZs in terms of local socioeconomic variables. Figure A.1 shows the distribution of population densities in the left panel and worker shares of the transportation machinery manufacturing (TTM) industry for each CZ in the right panel, using data from the 2007 ESS. In both panels, darker shades indicate that the relevant values are high. From the left panel, as one expects, the metropolitan areas facing the Pacific Ocean such as Tokyo, Osaka, Nagoya, Fukuoka have high population densities. In contrast, from the right panel, we may see that in terms of transportation machinery-manufacturing share, areas including Toyota City, Hamamatsu City, Kure City, and Kita-Kanto are standing out.[10] Also, note that in contrast to other labor market concepts such as UEAs, our methodology permits the analysis of the whole of Japan, providing a comprehensive and exhaustive analysis of all local employment areas throughout the country.

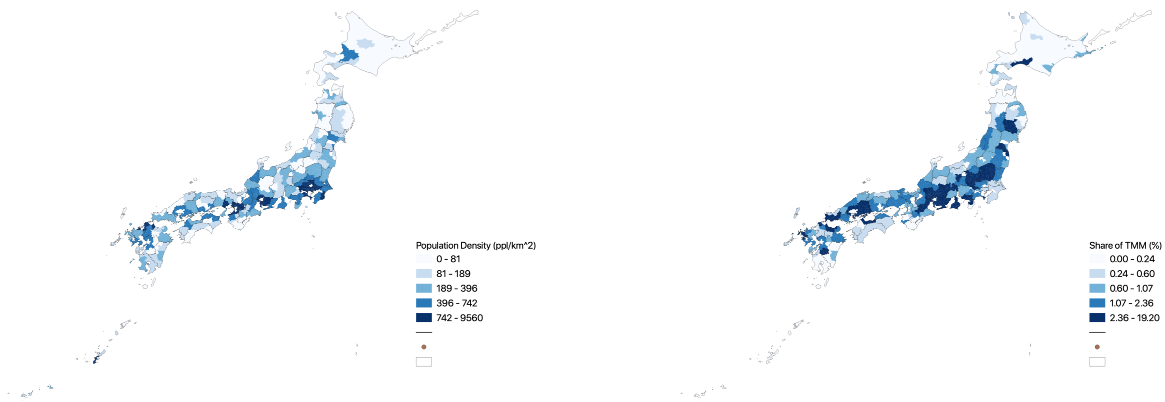
A.2 Relationship to the Railway Network

How are CZs related to commuting infrastructure? To address this question, we studied the overlap of our CZs and the railway network in Japan by obtaining ESRI shape files of railway networks from the National Land Numerical Information download service. We then layered the 2015 railway network on our map of CZs in 2015, resulting in Figure A.2. The figures show a pattern that the railways typically go through central areas of each CZ. This makes sense given that the railway is a major choice of commuting, particularly in the Kanto region of Japan. As classical commuting models indicate that central business districts agglomerate along railways (stations, in particular), around which workers commute (Alonso et al., 1964; Mills, 1967; Muth, 1969), it is natural that railways would cross along the interior of CZs rather than along the borders.

[10]Toyota City is the historic hometown to Toyota Motor Corporation, Hamamatsu City to Honda Motor Corporation, Suzuki Motor Corporation, and Yamaha Corporation, and Kure City is located near the headquarters of Mazda Motor Corporation. These cities also have many subcontracting companies who supply automobile parts and so are located close to these large automobile producers.

Figure A.1: Distribution of Socio-economic Characteristics

(a) Population Density (b) Share of Workers in the TTM Industry



Note: Figures created by authors based on Population Census data, with categorization by quintiles. The ESRI shape file for creating the map was generated by Kirimura et al. (2011). The "TTM" industry stands for the Transportation Machinery Manufacturing industry.

B Additional Figures and Tables

B.1 Delineation of Osaka and Nagoya Metropolitan Areas

Figure A.2 shows the 2015 CZ delineation with railroads for the Osaka and Nagoya metropolitan areas.

B.2 1980 CZs and UEAs

Figure B.2 shows the CZs and UEAs in the Kanto region in 1980.

B.3 Catchment Rate Analysis for Each Year

Figures B.3, B.4, B.5, and B.6 show the proportion of workers who commute within a geographic unit for each year.

C Guidance for Panel-Data Analysis

As econometric analysis using panel data is a quite effective tool in controlling for unobservable characteristics of the individual analysis unit, it has been used extensively in recent econometric studies. Since our method is not restricted to any specific years of data, it is possible to construct several alternative CZs depending on the choice of base year and unit codes. In our study, we simply provided all combinations of these for a broad choice set available for researchers. Here, we provide some brief guidance as to which set of CZ delineations should be used depending on the purpose of the analysis at hand.

First, note that for each year, we employed both *original* and *harmonized* municipality codes to construct CZs. The original municipality codes are those originally published by Japan's Ministry of Internal Affairs and Communications (MIC) at that time. However, since municipalities have merged and split over the years, the original municipality codes are not consistent across years. The harmonized municipality code, on the other hand, was generated with the 2015 municipality delineation using the HAC method described in Section 3.[11]

Given the definitions of the original and harmonized codes, we suggest the following guideline:

1. Use the original code when the researcher wishes to know the exact employment area for each period.

2. Use the harmonized code when the researcher wishes to analyze changes in employment areas, as the harmonized code facilitates a consistent comparison of CZs over time.

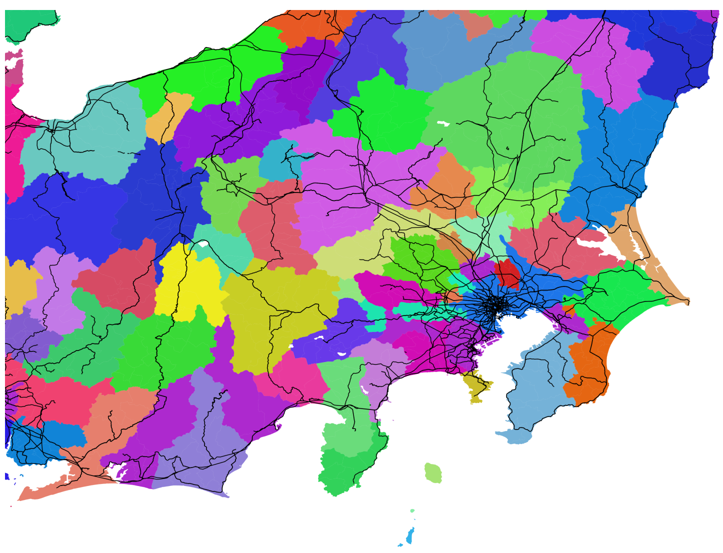
3. In either case, when analyzing with a fixed effect and CZ-year panel data, fix the analysis to one year, using a clear base year if any. For example, if the analysis period is 1990-2007, the researchers may use CZs based on either original or harmonized 1990 municipality code.

[11] We set the harmonization year as 2015 since the trend of the total number of municipalities is decreasing. Therefore, later years have fewer problems associated with splitting municipalities when creating harmonized codes.

Figure A.2: 2015 CZ Delineation with Railroads

(a) Nationwide

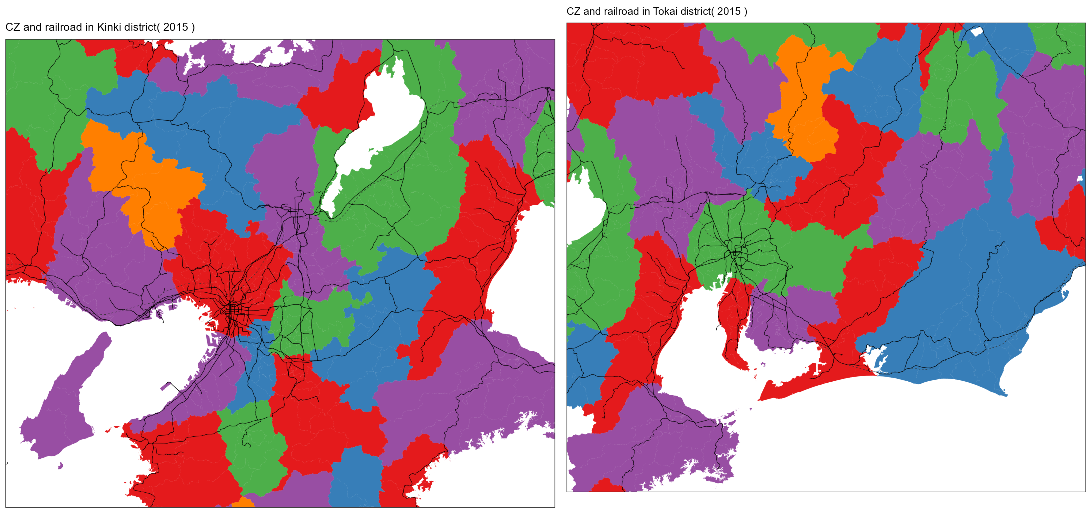
(b) Kanto district



Note: Map was created by authors based on the 2015 Population Census and hierarchical agglomerative clustering with average linkages, and harmonized municipality codes as an input to the algorithm. The railway network was sourced from Railways (line) files from the National Land Numerical Information download service. The ESRI shape file for creating the map was generated by Kirimura et al. (2011).

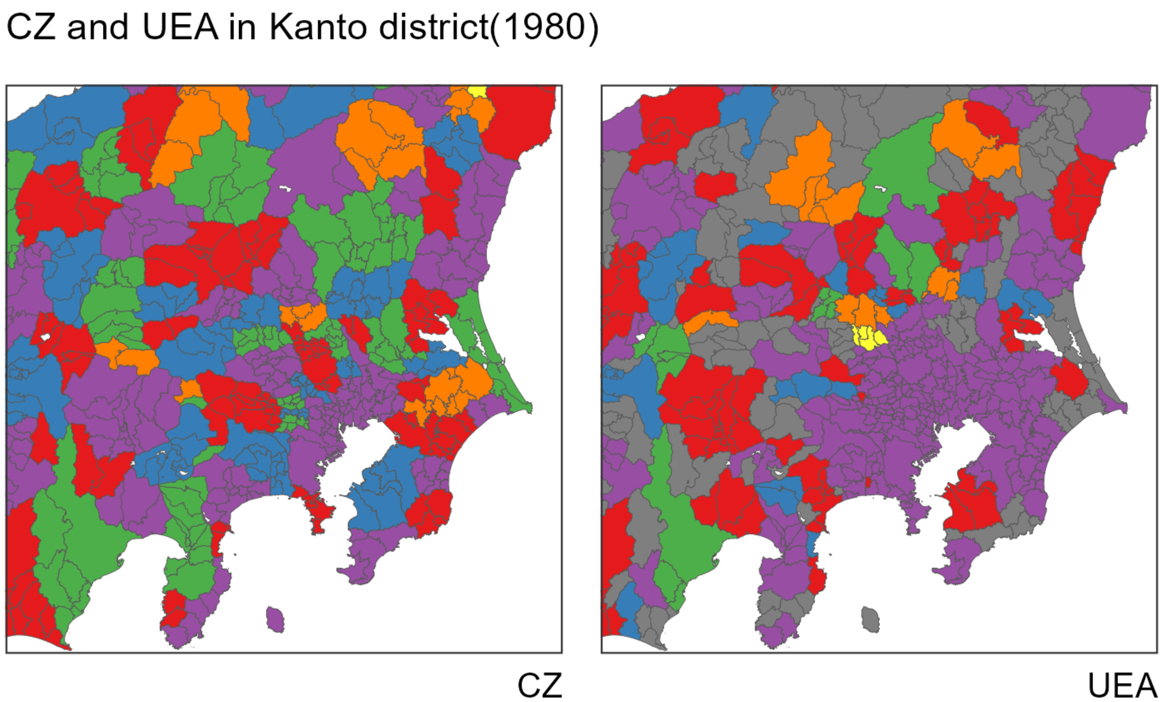
Figure B.1: 2015 CZ Delineation with Railroads

(a) Osaka Metropolitan Area (b) Nagoya Metropolitan Area



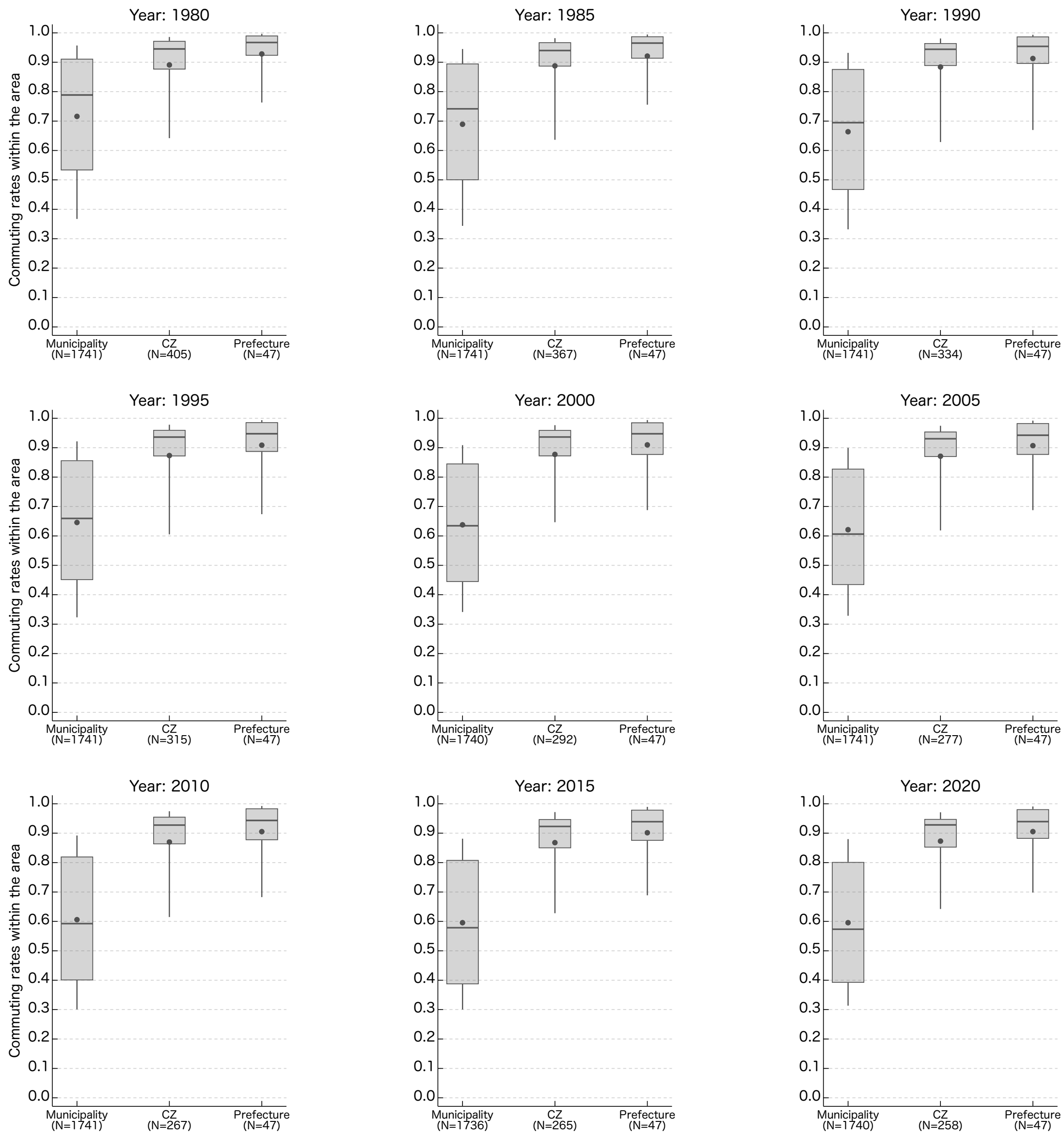
Note: The figure shows the commuting zones (CZs) overlaid with the railway network for the Osaka and Nagoya metropolitan areas. Railways typically run through the central areas of each CZ, reflecting their importance as commuting infrastructure.

Figure B.2: CZs and UEAs in Kanto region in 1980



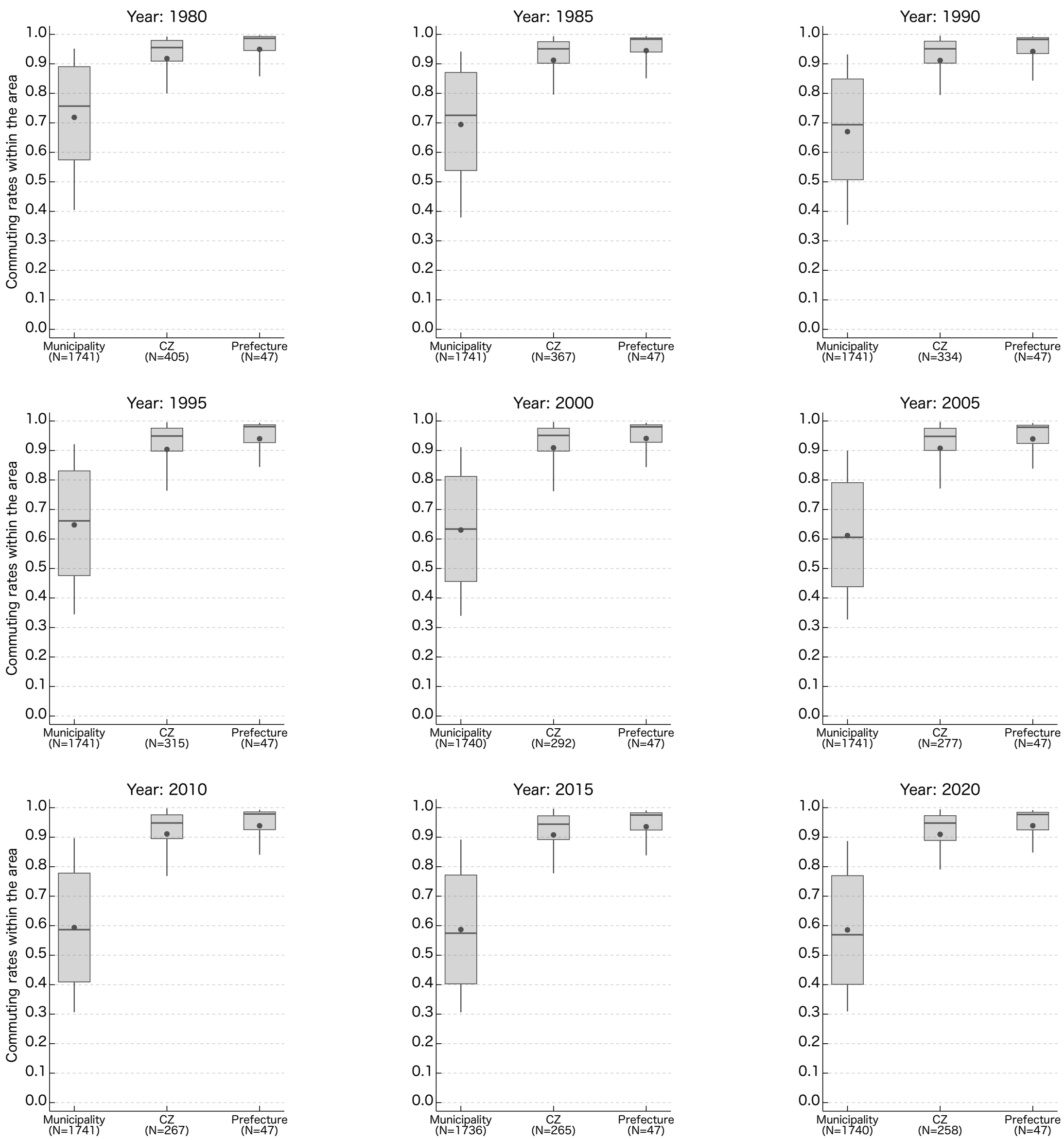
Note: The left panel shows the commuting zones (CZs) and the right panel shows the urban employment areas (UEAs). Gray color shows the area that is not included in UEAs.

Figure B.3: Proportion of Workers Who Commute within a Geographic Unit for Each Year: Population weighted and Harmonized CZs



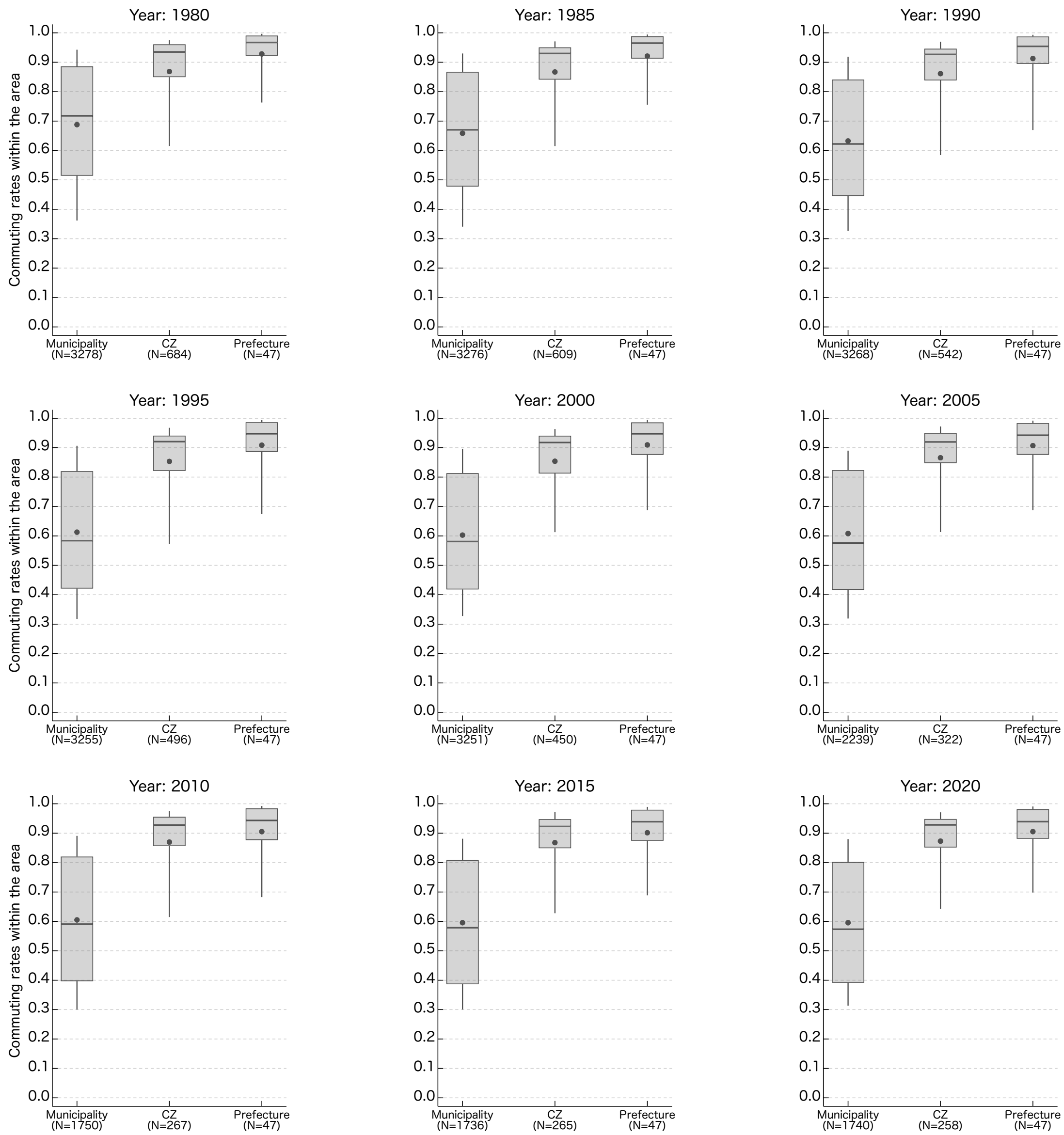
Note: This figure shows the distribution of the share of residents who commute within their own area, measured at the levels of prefectures, commuting zones, and municipalities. The distributions are calculated separately for each year from 1980 to 2020. For commuting zones, the harmonized CZs are used. In each case, the distribution is weighted by the population within each unit. The dot indicates the mean, the line in the middle of the box indicates the median, and the upper and lower ends of the box indicate the 75th and 25th percentiles. The line above and below the box show the 90th and 10th percentiles.

Figure B.4: Proportion of Workers Who Commute within a Geographic Unit for Each Year: Population unweighted and Harmonized CZs



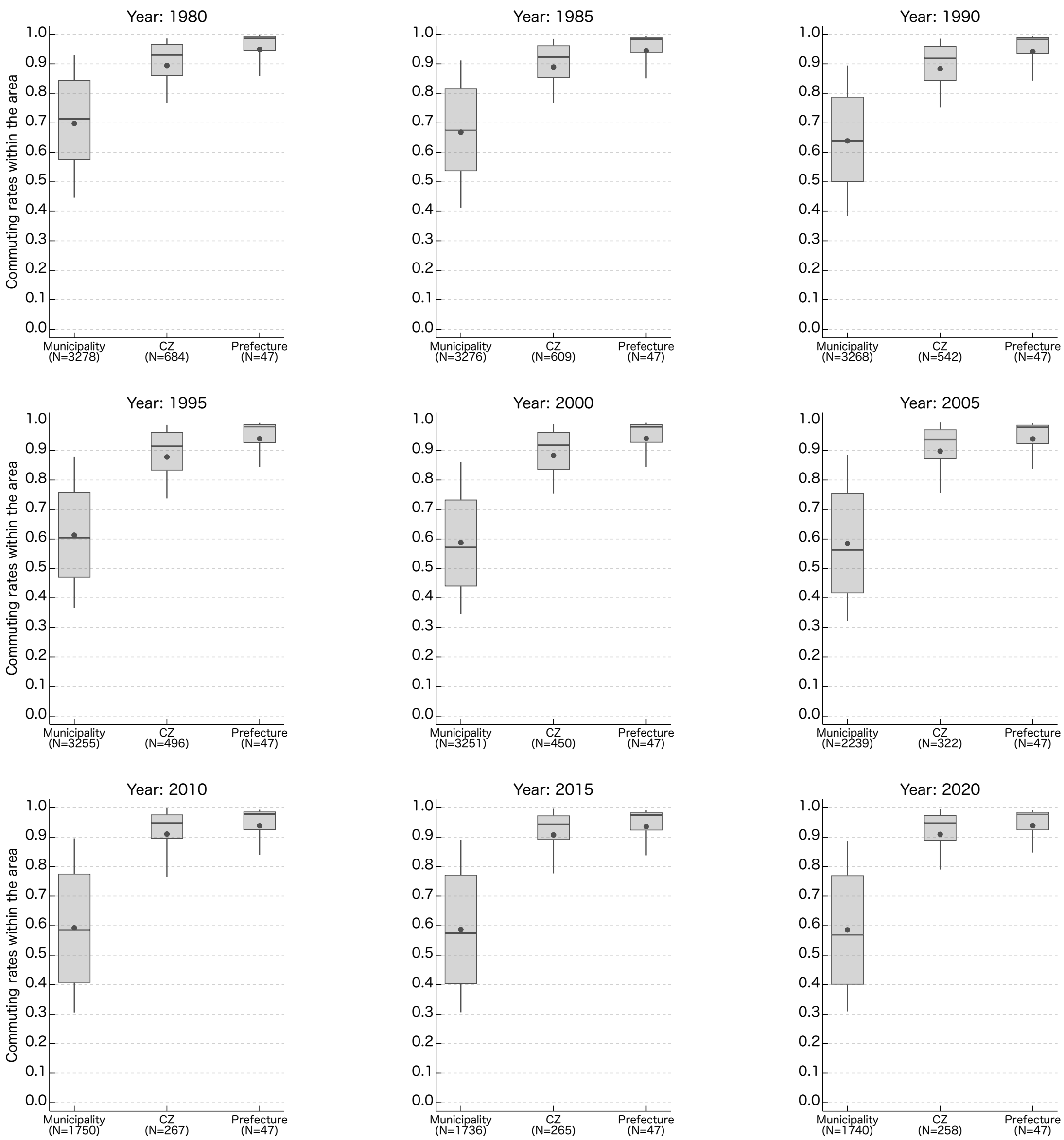
Note: This figure shows the distribution of the share of residents who commute within their own area, measured at the levels of prefectures, commuting zones, and municipalities. The distributions are calculated separately for each year from 1980 to 2020. For commuting zones, the harmonized CZs are used. The dot indicates the mean, the line in the middle of the box indicates the median, and the upper and lower ends of the box indicate the 75th and 25th percentiles. The line above and below the box show the 90th and 10th percentiles.

Figure B.5: Proportion of Workers Who Commute within a Geographic Unit for Each Year: Population weighted and Original CZs



Note: This figure shows the distribution of the share of residents who commute within their own area, measured at the levels of prefectures, commuting zones, and municipalities. The distributions are calculated separately for each year from 1980 to 2020. For commuting zones, the original CZs are used. In each case, the distribution is weighted by the population within each unit. The dot indicates the mean, the line in the middle of the box indicates the median, and the upper and lower ends of the box indicate the 75th and 25th percentiles. The line above and below the box show the 90th and 10th percentiles.

Figure B.6: Proportion of Workers Who Commute within a Geographic Unit for Each Year: Population unweighted and Original CZs



Note: This figure shows the distribution of the share of residents who commute within their own area, measured at the levels of prefectures, commuting zones, and municipalities. The distributions are calculated separately for each year from 1980 to 2020. For commuting zones, the original CZs are used. The dot indicates the mean, the line in the middle of the box indicates the median, and the upper and lower ends of the box indicate the 75th and 25th percentiles. The line above and below the box show the 90th and 10th percentiles.