Recalculating ... : How Uncertainty in Local Labor Market Definitions Affects Empirical Findings

Andrew Foote[1], Mark J. Kutzbach[2], and Lars Vilhuber[1,3]

[1]Center for Economic Studies, U.S. Census Bureau
[2]Federal Deposit Insurance Corporation
[3]Labor Dynamics Institute, Cornell University. Corresponding author: lars.vilhuber@cornell.edu

October 8, 2020

Abstract

This paper evaluates the use of commuting zones as a local labor market definition. We revisit Tolbert and Sizer (1996) and demonstrate the sensitivity of definitions to two features of the methodology: a cluster dissimilarity cutoff, or the count of clusters, and uncertainty in the input data. We show how these features impact empirical estimates using a standard application of commuting zones and an example from related literature. We conclude with advice to researchers on how to demonstrate the robustness of empirical findings to uncertainty in the definition of commuting zones.*

Keywords: Local labor markets; commuting; measurement error; Sensitivity analysis
JEL codes: C13 (Estimation) J21 (Labor force and Employment, Size, and Structure) C63 (Computational Technique)

*The analysis, conclusions, and opinions expressed herein are those of the author(s) alone and do not necessarily represent the views of the U.S. Census Bureau or the Federal Deposit Insurance Corporation. All results have been reviewed to ensure that no confidential information is disclosed, and no confidential data was used in this paper. Much of the work developing this paper occurred while Mark Kutzbach was an employee of the U.S. Census Bureau.

1 Introduction

Local labor markets are an important unit of analysis in economics and are a basis of many place-based public policies. Theoretical papers emphasize the characteristics of local labor markets including common wage and rent levels (Roback, 1982; Moretti, 2011) as well as common job-finding and unemployment rates (Head and Lloyd-Ellis, 2012; Schmutz and Sidibé, 2014). Theory often assumes fixed costs for transferring jobs or workers between labor markets, so imbalances can persist across labor markets for some time. For empirical analysis, researchers interested in estimating the effect of some local, exogenous shock on labor market outcomes must decide how to define the set of affected jobs or workers. Researchers examining labor markets in the United States often turn to one of several standard geographic definitions that are widely known and compatible with publicly available economic data, including: states (Blanchard and Katz, 1992; Wozniak, 2010; Kennan and Walker, 2011), metropolitan areas (Bound and Holzer, 2000; Card, 2001; Notowidigdo, 2020; Diamond, 2016), and counties (Monte, Redding and Rossi-Hansberg, 2018; Foote, Grosz and Stevens, 2019).

Another labor market definition preferred for some research topics is commuting zones, or clusters of locations with strong home-to-work commuting ties amongst one another. For the United States, commuting zones consist of counties aggregated together using hierarchical clustering of flows data (Tolbert and Sizer, 1996) (henceforth, TS). Commuting zones are similar to metropolitan areas in that they are meant to capture economic integration that does not necessarily conform to regional political boundaries, such as states (Office of Management and Budget, 2000, 2010). Unlike metropolitan areas, commuting zones have no urbanized area size requirements and span the entire United States, allowing researchers to measure effects for the entire country rather than just the set of metropolitan areas (or the complements of metropolitan areas within a state). Commuting zones have been used in a number of influential papers in the labor economics literature, including Autor, Dorn and Hanson (2013), Chetty et al. (2014), Amior and Manning (2018), Restrepo (2015), and Yagan (2016). Hierarchical clustering and other methods have been implemented with similar data in other countries both for official statistics and research purposes.[1]

Despite their widespread use, the methodology underlying commuting zone definitions and its impact on empirical estimates has not received much scrutiny, and researchers typically do not assess how sensitive their findings might be to commuting zone design, measurement, and uncertainty.

Our research contributes to the understanding of empirical analyses that use commuting zones and similarly defined areas. First, we demonstrate that commuting zone definitions are sensitive to both a dissimilarity threshold choice in the clustering process, which governs the number of clusters, and to uncertainty in the underlying worker flows data, which governs the strength of ties between pairs of counties. Our findings suggest that researchers should evaluate the sensitivity of their results when using commuting zones. Second, we show with two empirical examples how these methodological issues impact estimates that use U.S. commuting zones as a local labor market definition. We propose a methodology for evaluating robustness of results using analysis areas defined by flows data. The code underlying the

[1] For surveys and summaries of local labor market definitions and clustering methods, see Casado-Díaz and Coombes (2011) and Franconi, Ichim and D'Aló (2017). For a recent implementation of the TS method in Portugal, see Afonso and Venâncio (2016).

empirical example (Foote, Kutzbach and Vilhuber, 2020) can be applied to any such flow data, as long as measures of uncertainty are available.

Our analysis relates to the literature on the modifiable areal unit problem (MAUP), and more specifically, the concern that model estimates using spatial data may be sensitive to the shape or size of geographic analysis areas. Investigations of MAUP in different contexts have come to varied conclusions, for example, Fotheringham and Wong (1991) find that aggregation and shape variation can significantly alter results, while Briant, Combes and Lafourcade (2010) find that measurement and model specification issues are of first order importance relative to area size, with area shape being the least important. Empirical papers that use commuting zones rarely address either the issues of zone definition (e.g. by controlling for cross-zone commuting) or MAUP issues more broadly (e.g. by checking robustness to different spatial units).[2] We demonstrate the consequences for multivariate analysis of uncertainly in the definition of areas defined by commuting flows data. This paper goes beyond the MAUP studies cited above by examining not only the sensitivity of estimates to area shape and size, but also examining the role of assumptions in zone definition underlying variation in definitions.[3]

2 Commuting Zone Data and Methodology

The Economic Research Service (ERS), an agency under the U.S. Department of Agriculture for which commuting zones were originally developed, distributes definitions on its website.[4] Commuting Zones (CZ) are especially relevant for the economic analysis of rural areas, a focus of ERS, because they include all counties, not just urban counties.

To describe the CZ method, we rely on a graph-theoretic model of commuting flows, where each county is a node and edges measure the intensity of commuting between nodes. We review two design components: the dissimilarity matrix, which measures how "far" nodes are from one another, and the clustering method, which decides how nodes are assigned to groups.

2.1 Data

TS generate the commuting zones using flows from the 1990 Journey to Work (JTW) data (U.S. Census Bureau, 2017*a*), which tabulates the commuting information from the 1990

[2]Niedzielskia, Horner and Xiao (2013) investigate MAUP issues in commuting flows, but focus on sensitivity of summary measures (for excess commuting), rather than a multivariate analysis as with the studies cited above. Other studies also consider the optimal definition of traffic analysis zones, but that is a much smaller spatial scale and a different focus than analysis of local labor markets. Afonso and Venâncio (2016) report having checked for robustness of results with a lower average-linkage commuting flow threshold.

[3]Briant, Combes and Lafourcade (2010) evaluate an area defined by commuting flows, but for comparison, they use either arbitrary spatial units of the same size or larger spatial units defined using different criteria.

[4]ERS released commuting zone definitions based on 1980, 1990, and 2000 commuting data. All three definitions are available at http://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas.aspx as of 2020-09-30. For an analysis of the historical methodology, including the use of expert opinion, see Fowler, Rhubart and Jensen (2016).

Census long-form (U.S. Census Bureau, 1992*b*).[5] County-to-county flows were estimated for 3,141 county equivalents for persons age 16 and older who reported being employed in the week prior to April 1, 1990.[6]

2.2 Methodology

The dissimilarity matrix, d, is a representation of the relative distance between all pairs of N counties. One can calculate d, where an entry d_{ij} is the dissimilarity of county i from county j, as follows:

$$d_{ij} = 1 - \frac{f_{ij} + f_{ji}}{\min(rlf_i, rlf_j)} \tag{1}$$

In the above equation, f_{ij} is the flow of commuters who live in county i and work in county j and f_{ji} is the opposite flow. The resident labor force in county i is $rlf_i = \sum_{j=1}^{N} f_{ij}$ (including f_{ii}), with a corresponding calculation for county j. Normalizing flows with the minimum resident labor force of a pair upweights the association of outlying areas with metropolitan cores. Note that disimilarity is symmetric, so $d_{ij} = d_{ji}$.

After constructing this dissimilarity matrix, TS use it as an input into their clustering method, the average-linkage hierarchical clustering algorithm (PROC CLUSTER in SAS software).[7] At the cluster level, the dissimilarity for a pair of clusters, C_K and C_L, is calculated as the average of dissimilarity among all pairs of counties between the clusters, written:

$$D_{KL} = \frac{1}{N_K N_L} \sum_{i \in C_K} \sum_{j \in C_L} d_{ij}, \tag{2}$$

where d_{ij} is calculated as in Equation 1 and N_K and N_L are the count of nodes, or counties in this case, in each cluster.

The hierarchical clustering method uses the dissimilarity matrix in the following way. To begin, every county is its own cluster, so $d_{ij} = D_{KL}$. The algorithm finds the lowest value of dissimilarity and combines those two counties together, forming a new cluster. It then recalculates a dissimilarity value between the new cluster and all other clusters. The process continues joining clusters (including singletons with only one county) with the lowest dissimilarity D_{KL} until all nodes are clustered. The designer may set a stopping threshold by choosing a maximum "cutoff height" H. If $D_{KL} > H$ then K and L do not merge.[8]

[5] Journey to Work data on county to county commuting flows are available for the 1990 Census, the 2000 Census, and 5-year samples of the American Community Survey at https://www.census.gov/hhes/commuting/data/commutingflows.html.

[6] Employment status is based on responses to the question "Did this person work at any time LAST WEEK?" Place of work is geocoded using the response to "At what location did this person work LAST WEEK?" Residence location is compiled from the mailing frame of the Census.

[7] The hierarchical clustering for this paper using PROC CLUSTER was generated using SAS software, Version 9.2 of the SAS System for Unix. Copyright ©2009 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

[8] To visually illustrate how the clustering algorithm works, Appendix Figure 9 displays the results of the clustering procedure for different values of H, focusing just on counties in California. In the top left-hand corner, at a height of $H = 0.80$, only a few counties have joined. As we increase the height to 0.88 and to 0.96, more counties are joined together. Finally, at a height of 1, almost all the counties have merged together,

TS processed the 1990 JTW data with this procedure, setting H to 0.98. We refer to the resulting set of commuting zones for 1990, published in Tolbert and Sizer (1996) and available today as Economic Research Service (2012), as "TS1990".

2.3 Our Replication

To assign each county to a unique cluster, TS processed the 1990 JTW data with the procedure outlined above. They set cluster height, H, to 0.98, divided the country into six overlapping regions, and performed the clustering algorithm separately in each region (due to computational constraints). Afterwards, they manually resolved conflicts in overlapping regions. These decisions have consequences for users. First, while the height cutoffs across regions are the same, a step in the algorithm that normalizes distances to unit mean will operate differently across regions, re-normalizing the dissimilarity matrix for each region. As a result, zones in the Northeast have different connectivity than zones in the South. Second, even with overlap allowances, some clusters may be constrained. Third, overlaps induce some undocumented subjectivity, since conflicts in cluster assignment for counties in the overlapping regions are inevitable. Eighteen states fall within overlapping regions and Arkansas is in four regions.

Rather than follow their methodology of dividing the country into regions, which required a subjective expert review, we run the hierarchical clustering algorithm on the entire country (we exclude Alaska and Hawaii). We choose a height cutoff, 0.9385 (compared to 0.98), that most closely replicates their original zones in terms of size distribution, though it results in 755 zones (compared to 741).[9]

The TS1990 zones and our replication thereof (which we will refer to as FKV1990) are depicted in Figure 1. Table 1 provides summary statistics. As would be expected given the difference in cluster count, the allocation of counties to clusters varies across the definitions. Our replication has more evenly sized clusters with fewer clusters made up of a single county. There are some notable large clusters in our replication, mostly in sparse areas (e.g. Montana; Northern California and Eastern Oregon).[10] These large clusters serve to underline the main weakness of the hierarchical clustering methodology - counties are forced to merge eventually, even if the links are relatively weak.[11]

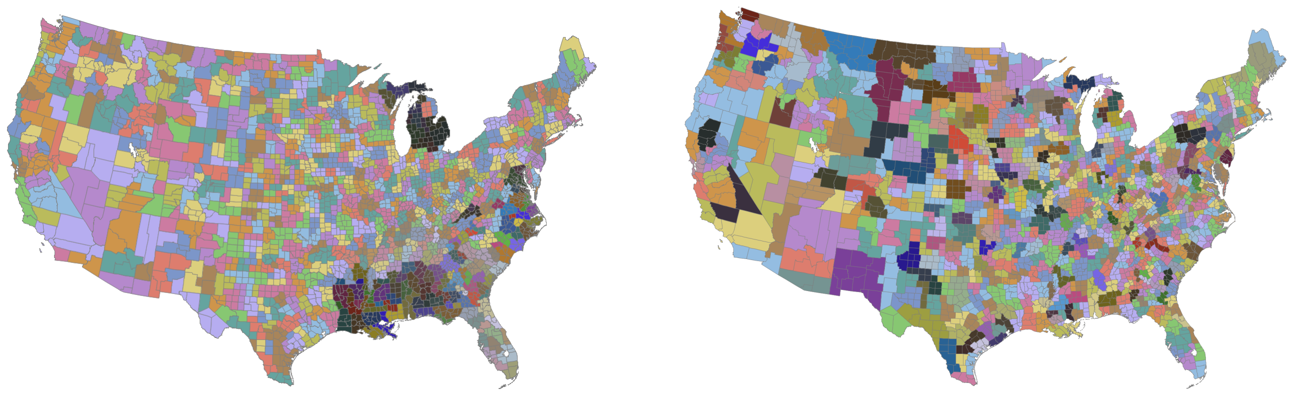
forming one large cluster and a few much smaller clusters.

[9]We find that the clustering algorithm, when attempting to produce the same cluster count as TS1990 with a national run, retains a residual cluster that spreads across many states. Only with the lower cluster height, and more clusters, does the residual cluster break up. We leave evaluation of alternate clustering methods for future work, with the emphasis in the present analysis on the sensitivity of estimates to zone definitions. We also note that this residual cluster is a feature of the hierarchical procedure, which selects new clusters based only on their incremental optimality. We show an example of such a map, using a cutoff of 0.945, in Figure 10 in the Appendix.

[10]It should be noted that TS1990 also has some large zones, including some in Nevada, Idaho, and California.

[11]One leading alternative clustering method is spectral clustering, in which a user specifies a number of clusters desired, and the algorithm minimizes total within-cluster distance.

Figure 1: Replication of Commuting Zones from TS: County Mapping



Commuting Zones - TS1990 **Replication of Commuting Zones - FKV1990**

Notes: In TS1990, the clustering procedure was done on six overlapping regions and then reconciled; in FKV1990 (author's calculations using the 1990 Census Journey to Work Tables), we performed the clustering procedure nationally. Full-size maps are available in the online appendix. See text for details.

Table 1: Replication of TS1990 Commuting Zones: Summary Statistics

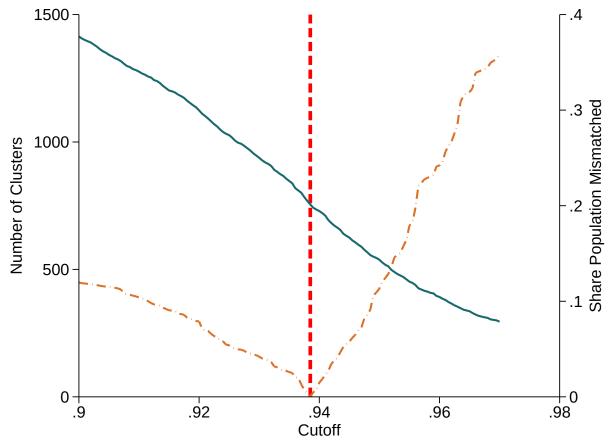
	TS1990	FKV1990
Mean Cluster Size	4.24	4.16
Median Cluster Size	4	4
SD Cluster Size	2.50	1.62
Number of Clusters	741	755
Number of Singletons	62	12
Cross-Commuting as Share of Population	0.060	0.107
Share of Population Mis-matched	0.17	

Notes: Both TS1990 and FKV1990 are based on Journey to Work tabulations from the 1990 Census. Summary statistics for TS1990 are from Table 8 of TS.

3 Design Sensitivity

While commuting zones are used by researchers as a convenient measure of local labor markets, they have a number of shortcomings for empirical research that are not regularly discussed in the literature. In this section, we evaluate the sensitivity of commuting zone definitions, focusing on two aspects of the TS methodology. First, we show that there is substantial ambiguity over the choice of when to stop merging the clusters, and that small changes in the chosen cutoff height can affect the number and size of the clusters. Second, we show that if there is uncertainty in the input data, the resulting commuting zone definitions can vary substantially. Overall, this uncertainty and subjectivity in the commuting zone definitions contributes to conventional standard errors understating the true level of uncertainty in the estimates, which we show when we return to this issue with our empirical replication in the next section.

Figure 2: Effect of Cluster Height on Number of Clusters



Note: Authors' calculations using methodology described in Section 2. A higher cutoff results in fewer clusters. The solid line is the number of clusters at a given cutoff and uses the left axis; the dashed-and-dotted line is the share of counties in a different commuting zone from our replicated zones, FKV1990, and uses the right axis. The red-dashed vertical line indicates the cutoff value for our replication, H^*.

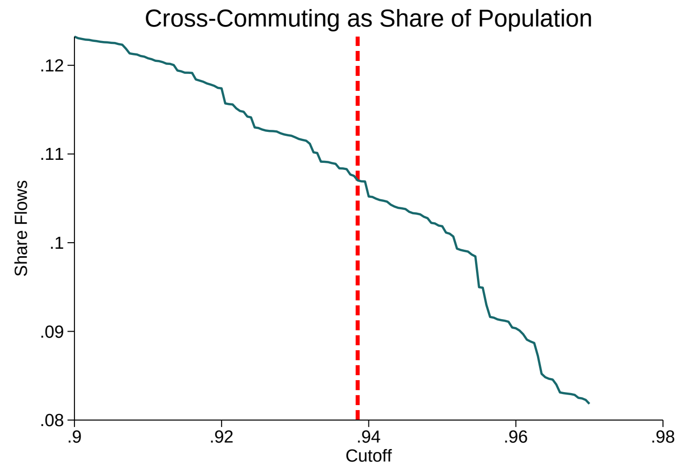
3.1 Choosing Cluster Height

One sensitive feature of the methodology used by TS is choosing the cutoff value above which no clusters can form, or the maximum dissimilarity, which determines the number of clusters. Tolbert and Killian (1987) describe the algorithm for choosing a cutoff value as follows (see page 15): "As a rule of thumb, a normalized average distance of 0.98 was considered sufficient distance between sets of counties to treat them as separate [Labor Market Areas]." The article does not provide an analysis of the sensitivity to changing the cutoff marginally up or down. TS, in an effort to minimize methodological differences between commuting zones for 1980 and 1990, use the same cutoff with no further evaluation for the 1990 data. In this subsection, we investigate how sensitive the clusters for our replication are to the choice of the cutoff value.

Figure 2 shows the number of clusters that form at various height cutoffs using the 1990 flows data, with the vertical line indicating the cutoff value we chose to replicate TS1990, $H^* = 0.9385$. As the cutoff rises, the number of clusters falls. Increasing or decreasing the cutoff has implications for the number of resulting clusters. Decreasing H to 0.9375 increases the number of clusters by 29, while using a cutoff of 0.93995 causes the number of clusters to decrease by 20. We also plot the share of the population that is in a different cluster than the replicated clusters, to illustrate how assignment changes further away from the cutoff. As we noted earlier, one particular issue with the methodology is that for cutoffs to the right of the vertical line, a large residual cluster forms, which contributes to the steeper slope in the share of the population mismatched.

Another way to measure the appropriateness of the cluster height is to measure the share of commuting flows that occur between clusters. If the clusters approximate local labor markets, those values should be lower, indicating internal cohesion and minimal overlap with

Figure 3: Cluster Height and Share Workers Commuting Between Clusters



Note: Authors' calculations using methodology described in Section 2. The solid line shows how the across-clusters average of the share of workers commuting outside of the cluster where they reside varies with the cluster height cutoff (greater height corresponds with fewer clusters). The red-dashed vertical line indicates the cutoff value for our replication, H^*.

outside areas (though if clusters are too large, they may encompass unrelated labor markets). Using the 1990 JTW data, we compute the share of total commuting flows whose destination is outside the cluster, average these values across all clusters, and plot these values by height cutoff in Figure 3. As the cutoff increases, the clusters get larger and there is less cross-cluster commuting. At the upper limit of the cutoff, with very few clusters, the value will tend towards zero.

This figure, taken together with Figure 2, shows some of the features of the cutoff decision. There is no discontinuity and no empirical guidance or broad consensus in the theoretical literature.[12] TS do not discuss their criteria for choosing these cutoffs.[13]

3.2 Sensitivity of Clustering Results to Underlying Error

Next, we analyze the extent to which the outputs of the TS methodology are sensitive to errors in the commuting flows data used for clustering. First, recall Equation 1 for the entries of the dissimilarity matrix. If f_{ij} is measured without error, then the distance between counties i and j is also measured without error. However, if the flows are measured with error, ϵ_{ij}, then observed flows may be defined as $\hat{f}_{ij} = f_{ij} + \epsilon_{ij}$. In this way, our observed $r\hat{l}f_i$, $r\hat{l}f_j$, and $\hat{d}_{ij}$ would also include error. We show this expression below (assuming without loss of generality that $rlf_i < rlf_j$):

[12] Decisions on clustering methods, clustering counts, and validation criteria depend on the application and are inherently somewhat subjective. Because clustering is an unsupervised method, there may be no indication of the ideal number of clusters (Halkidi, Batistakis and Vazirgiannis, 2001).

[13] Other designers have made different sizing choices. For example, Employment Areas in France, also based on commuting flows, are much smaller and would be equivalent to splitting the United States into over 4,700 units (Briant, Combes and Lafourcade, 2010)

$$
\begin{aligned}
\hat{d}_{ij} &= 1 - \frac{\hat{f}_{ij} + \hat{f}_{ji}}{r\hat{l}f_i} \\
&= 1 - \frac{f_{ij} + f_{ji}}{rlf_i + \sum_j \epsilon_{ij}} + \frac{\epsilon_{ij} + \epsilon_{ji}}{rlf_i + \sum_j \epsilon_{ij}}
\end{aligned}
\tag{3}
$$

Even if $E[\epsilon_{ij}] = 0$, that does not imply that $E[\hat{d}_{ij}] = d_{ij}$. Furthermore, we cannot rely on the limit properties of the error distribution, because we only have one realization of the commuting flow, which is calculated from survey responses. Additionally, we know that ϵ_{ij}/f_{ij} is larger for small flows. Errors will increase d_{ij} for some small counties and decrease it for others. Because of the hierarchical nature of the clustering method, this error will affect the formation of all other clusters in the data.[14]

To demonstrate how this measurement error affects the outcome of the clustering procedure, we project the published margins of error (MOE) from the 2009-2013 ACS Journey to Work data (U.S. Census Bureau, 2017*b*) onto the 1990 Journey to Work data (for which no uncertainty measures were published).[15] We resample these MOEs to obtain different realizations of the commuting zones in the following way:

1. For each origin-destination pair (i, j), we draw ϵ_{ij} from a normal distribution with mean 0 and standard deviation $MOE_{ij}/(1.645)$, since the MOE is scaled to be the 90% confidence interval.[16]

2. Calculate the new flow value, $\hat{f}_{ij} = f_{ij} + \epsilon_{ij}$, with negative values set to zero.[17]

3. Re-calculate each dissimilarity matrix entry $\hat{d}_{ij}$.

4. Re-run the hierarchical clustering procedure, still using the FKV1990 cutoff.

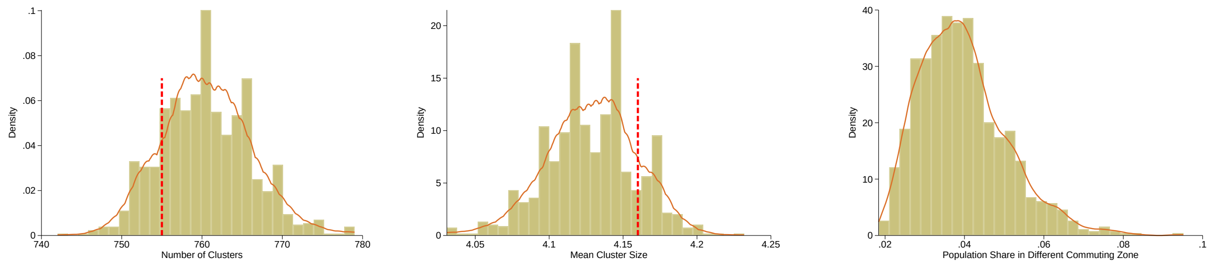
[14] In addition, because heights are normalized in the procedure, the dissimilarities of all clusters will change relative to the cutoff.

[15] To project MOEs from the ACS to 1990 flows, we calculate a ratio, MOE_{ij}/f_{ij}, representing the degree of uncertainty for a flow in the ACS. To reflect the range of MOEs across similarly sized flows, we calculate the mean and standard deviation of these ratios within flow size bins, defined by 1990 flow percentile, of: 0-50; 50-90; 90-95; 95-99; and 99+ (see Appendix Table 4). For each 1990 flow, we draw from the distribution of ratios calculated with the ACS in the corresponding bin. Note that the Census long-form is designed to be a one-in-six sample for one year, while the ACS data covers 5 years with a one-in-fifty sample each year. The smaller sample size of ACS typically results in higher margins of error for comparable statistics. The uncertainty implied by our implementation likely overstates the underlying MOEs in the 1990 flows. For more information on the construction of the ACS MOEs, see U.S. Census Bureau (2014, pages 10-12).

[16] In doing this, we assume that $\epsilon_{ij} \perp \epsilon_{ik} \forall k$, for simplicity. In reality, it is likely that $corr(\epsilon_{ij}, \epsilon_{ik}) < 0$, which means in our setting that we are understating the error by treating them as independent. In the Journey to Work data, there are likely some origin-destination pairs that are not reported due to the sample design. In our current resampling approach, we only resample from non-zero flows in the data. A more complete approach could model the likelihood that a zero reported is actually a positive flow, and resample accordingly, but that is beyond the scope of this paper. For more detail on the 1990 Decennial Census sample design, consult U.S. Census Bureau (1992*a*).

[17] 65 percent of flows are not statistically different from zero and are at risk to be censored, but these tend to be small flows and account for only 1.7 percent of jobs.

Figure 4: Results from Re-sampling Commuting Flows



(a) Number of clusters (b) Average Number of Counties (c) Share of Mismatched Population

Notes: Results from the resampling procedure described in Section 3.2. In subfigures (a) and (b), the red vertical line corresponds to the value from our benchmark commuting zones, FKV1990, as described in Section 2. Subfigure (c) gives the distribution of mismatch from resampling compared with FKV1990.

5. Store the re-sampled clusters for sensitivity analysis.

We iterate over this procedure 1,000 times in order to obtain distributions for these statistics, which are shown in Figure 4, where the red-dashed vertical lines are the values for our replication, FKV1990, obtained using the published 1990 data. The figures show that the average cluster size varies considerably from the result that the published figures would yield and that the realized outcome is not near the median for the number of clusters or mean cluster size. Additionally, the median share of the population that is mismatched is about 4% of the US population.[18]

4 Empirical Sensitivity

In the previous section, we showed that there are two margins on which clustering methodologies are sensitive: uncertainty in the input data and the choice of the number of clusters. However, these issues are only important for applied research to the extent that the uncertainty impacts empirical estimates in a significant way. In this section, we assess the impact of the uncertainty in commuting zones for two specific empirical examples. First, we estimate a model relating changes in local labor demand on receipts of unemployment insurance in the same local area, and assess the sensitivity of the estimated parameter to variations in the definitions of the local area. Second, we conduct a replication of Autor, Dorn and Hanson (2013) on the effect of trade competition on job loss, and assess the sensitivity of the estimated parameters, and their conclusions, to variations in the definition of commuting zones.

[18]To get a sense for how robust these re-sampling results are over time, we replicate the analysis reported in Figure 4 using the 2009-2013 ACS data. This has the additional advantage of using the MOEs specifically constructed for the data, rather than projecting them onto the 1990 commuting flows. The results are shown in Figure 13 in the Appendix. Mismatched population is actually higher, with a median of almost 5% of the population and a large right tail. The distributions of clusters and counties (panels a and b) are tighter when using the JTW 2009-2013 data, but the average number of clusters (panel a) is smaller and the mean cluster size (panel b) is higher than the values in Figure 4, which reflects higher integration between counties during this time period.

4.1 Case Study 1: Unemployment Receipt and Labor Demand

We measure labor demand in 1990 commuting zones using a standard Bartik (1991) measure:[19]

$$Demand_{K,t} = \sum_{s} \frac{Emp_{Ks,1990}}{Emp_{K,1990}} (log(Emp_{s,t}) - log(Emp_{s,t-1})) \tag{4}$$

Equation (4) states that demand for area, or cluster, K in year t is a weighted sum over all industry sectors s, where the weights are the original (1990) industry shares of employment in an area. National changes in employment for an industry, denoted $(log(Emp_{s,t}) - log(Emp_{s,t-1}))$, affect areas according to the original intensity of employment in each area.

To measure the components of Equation (4), we use the Bureau of Labor Statistics' Quarterly Census of Employment and Wages for 1990-2015 (Bureau of Labor Statistics, 2020) to compute average annual employment for each of the twenty NAICS industry sectors. We measure the total receipts of unemployment insurance using the Bureau of Economic Analysis's Regional Economic Accounts data (Bureau of Economic Analysis, 2019).

Our estimating equation is

$$log(UIReceipts)_{K,t} = \alpha Demand_{K,t-1} + \gamma_K + \delta_t + \epsilon_{K,t} \tag{5}$$

where $Demand_{Kt}$ is measured according to (4), and γ_K and δ_t are area and year fixed effects. We first estimate (5) for the commuting zone definitions TS1990 and FKV1990, with 25 years of data across each set of commuting zones. We then present estimates using the various realizations of commuting zones based on the results from Section 3.

4.1.1 Results of Estimating Equation

Table 2: Effect of Labor Demand on Unemployment Receipt

	Using TS1990 CZs	Using FKV1990 CZs
$Demand_{it}$	-7.0163	-8.3953
	(0.9660)	(0.7419)
N	18441	18700
R-squared	0.984	0.978

Notes: Table from author's calculations, based on Equation 5. Column 1 uses TS1990 commuting zone definitions, Column 2 uses FKV1990 commuting zones as described in Section 3. Standard errors in parentheses are clustered at the commuting zone level. All coefficients are significant with p-values less than 0.01.

Results in Table 2 show that when labor demand increases, unemployment receipts fall.[20] To scale our results, a one standard deviation change in labor demand (0.02) causes

[19] Bound and Holzer (2000), Notowidigdo (2020) and Autor, Dorn and Hanson (2013) use this measure of labor demand; the latter two papers estimate the effects of demand shocks on uptake of public assistance.

[20] Observations approximately equal the product of 25 years by the number of clusters in each definition. A few observations are missing due to missingness in certain states for the early period in the BEA data.

Figure 5: Differences in Effect Based on Cluster Cutoff

Note: Author's estimates of α in Equation (5) while varying the cluster height cutoff. The solid red line gives the point estimate of α for a given cluster definitions at each height, the gray area denotes the 90 percent confidence interval. The red dashed line is the reference point estimate for clusters using the benchmark definition FKV1990 at $H^* = 0.9385$.

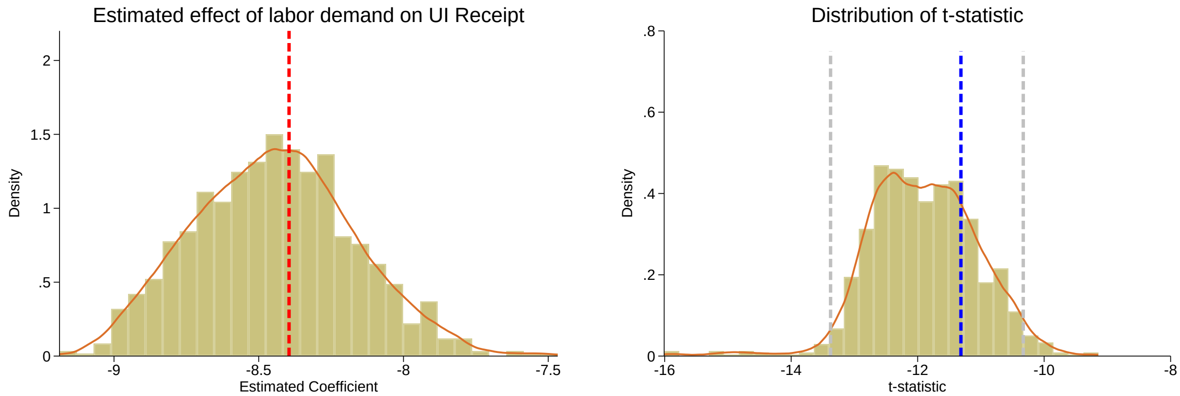
unemployment receipts to fall by approximately 14-16 percent. The magnitude of the point estimate varies depending on which set of commuting zone definitions are used but the difference is not statistically significant.

4.1.2 Sensitivity to Chosen Cutoff

To demonstrate how the cutoff choice affects estimates of α from Equation (5), we use the clusters generated based on cutoffs between 0.9 and 0.97. The count of clusters in this range falls from 1,423 to 303, so observations drop and the standard errors rise with the cluster height cutoff. Again, we note that for values of H slightly larger than $H^* = 0.9385$, a large cluster forms that is not consistent with our interpretation of local labor markets. For this reason, we take the estimates for cutoff values below H^* to be more reasonable.

Figure 5 shows that the estimated magnitude of $\hat{\alpha}$ increases as the cutoff height H increases (the number of clusters declines). With a greater number of clusters at lower cutoffs, commuting zones may be underbounded, or not large enough to capture the relevant local labor market for resident workers (see Figure 3). Variables describing commuting zones would then be measured with error, which would lead to attenuated parameter estimates. As clusters are consolidated, the magnitude of the effect increases, but then levels off.

Figure 6: Distribution based on Realizations of CZs



(a) Coefficient

(b) T-Statistic

Note: Histogram plots estimates of β from Equation 5, based on commuting zone realizations as outlined in Section 3. Red vertical dashed line in Panel (a) shows estimate $\hat{\beta}$ for FKV1990 commuting zones. Histogram plots t-statistics derived from estimating Equation (5), based on commuting zone realizations as outlined in Section 3. In Panel (b), blue vertical dashed line is t-statistic using FKV1990, and gray vertical lines are the 2.5th and 97.5th percentiles of the t-statistic distribution.

4.1.3 Sensitivity to Errors in Flows Data

To investigate how sensitive the results are to different commuting definitions, we re-estimate (5) using 1000 realizations of commuting zones (see Section 3.2).

The distribution of $\hat{\alpha}$ from this exercise is plotted in Figure 6 panel (a). The red vertical line indicates the value $\hat{\alpha}(FKV1990) = -8.3953$ obtained with the FKV1990 CZs. In this particular example, the distribution is symmetric with the coefficient centered on the distribution, but as the next example shows, that is not always the case.

Another way to summarize the results of this exercise is to look at the distribution of t-statistics, which incorporates information about the standard error of $\hat{\alpha}$ into the analysis as well, and comparing that distribution to the critical values.

Figure 6 panel (b) shows the distribution of t-statistics obtained from estimating Equation (5). The blue vertical dashed line is the original t-statistic, and the light gray vertical lines are the 2.5th and 97.5th percentiles. Clearly, in this application the result is still significant, because the entire confidence interval of t-statistics is less than the critical value (-1.96).

Importantly, the estimated t-statistic from the original regression (blue vertical line) is not the median of the distribution, which warrants caution on the part of the researcher, because the zones that are used by the researcher may represent an outlier in terms of statistical significance. This exercise demonstrates that there is additional uncertainty induced by the construction of the commuting zones that is not addressed in empirical estimates that use these definitions, which may overstate the precision of results.

4.2 Case Study 2: The China Syndrome

Autor, Dorn and Hanson (2013) estimate the impact that increased trade competition from

China had on manufacturing employment in the United States. The identifying variation used in their analysis is differences in industrial composition across (TS1990) commuting zones, which varies the exposure of workers to import penetration following China's entry into the World Trade Organization in 2001. We replicate the analysis, assessing the robustness of the result to variations in the definition of the commuting zones. While we find that the overall result is robust, we show that the magnitude of employment effects varies substantially and non-monotonically with the number of clusters.[21]

The main estimating equation in Autor, Dorn and Hanson (2013) is:

$$\Delta L_{it}^{m} = \gamma_t + \beta_1 \Delta IPW_{uit} + \beta_2 X_{it} + \epsilon_{it} \tag{6}$$

where ΔL_{it}^{m} is the decadal change in manufacturing employment in commuting zone i following year t, ΔIPW_{uit} is the import exposure growth measure, and X_{it} are control variables. All regressions are weighted by population share of the commuting zone.

4.2.1 Replication Results

Since we use slightly different data sources to measure manufacturing employment, we first establish a baseline replication, comparting to the main estimates from Autor, Dorn and Hanson (2013, their Table 2). Table 3 shows a comparison of original headline estimates to variations in data sources in our replication.[22] Each cell in the table is a coefficient from a different regression. For simplicity, we show estimates for the time period 1990-2000 only. The first column shows estimates from Autor, Dorn and Hanson (2013, Table 2, Column 1). Columns 2-5 show estimates of the same model, replacing the import exposure measures (Column 2), the estimates of the change in manufacturing employment and the weights (Column 3), and the clustering variable (TS1990 instead of state, Column 4, and FKV1990, Column 5).

In Autor, Dorn and Hanson (2013), the interquartile range of import exposure growth is approximately \$1000. Their estimates (in column 1) imply a reduction in manufacturing employment of 0.88 percentage points, while our estimates imply a decrease of 0.87 percentage points. The difference is not statistically significant. Overall, the estimates are robust, giving us confidence that we are properly replicating their initial findings. We now turn to demonstrating how these estimates are affected by the concerns with the commuting zone definitions themselves.

4.2.2 Sensitivity to Chosen Cutoff

In Section 3.1 we showed that the decision of where to stop the clustering process was rather arbitrary, since there is no clear guidance on what cutoff is most appropriate. To demonstrate how the cutoff choice affects the estimate of β_1 from Equation 6, we generate clusters based on cutoffs between 0.9 and 0.97 and estimate the model using the resulting clusters.

[21] We want to acknowledge that Autor, Dorn and Hanson were incredibly helpful in the process of replicating their paper, both in providing data and helping to troubleshoot, as well as being receptive to this exercise.

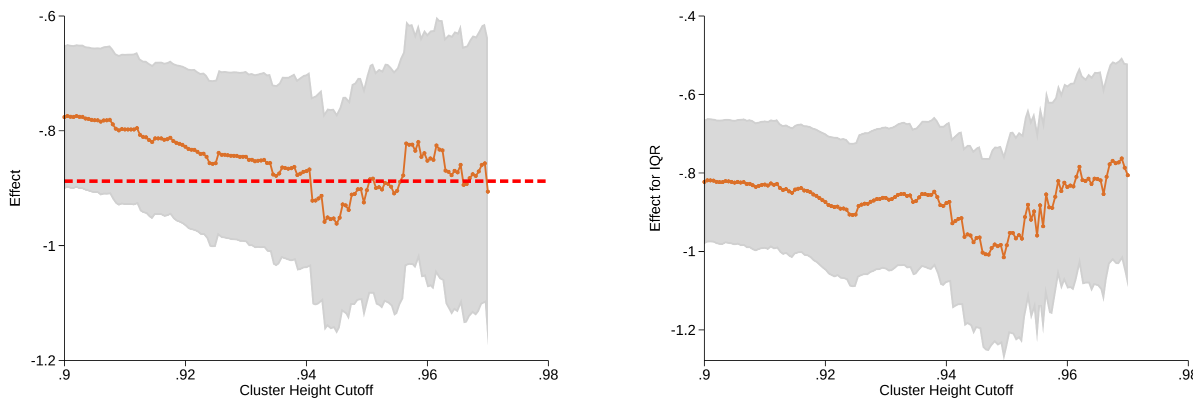
[22] We use county level manufacturing employment from the decennial census (Minnesota Population Center, 2016), while Autor, Dorn and Hanson (2013) map the PUMS data at the PUMA level into the commuting zone using David Dorn's crosswalk.

Table 3: China Syndrome Replication and Comparison, 1990-2000

	ADH Estimate	Our RHS	Our LHS and Weight	CZ Clustering	Using FVK1990
$\Delta IPW_{cz,t}$	-0.8875	-0.8871	-0.8748	-0.8748	-0.8725
	(0.1812)	(0.1811)	(0.1527)	(0.1243)	(0.0628)

Notes: Table from author's calculations, using data from Autor, Dorn and Hanson (2013) and constructed data, based on Equation 6. Column 1 is Table 2, Column 1 from ADH (2013). Column 2 replaces their measure of import exposure with ours. Column 3 replaces their measure of change in manufacturing employment and CZ-specific weights with ours. Columns 4 and 5 cluster by commuting zone, rather than state. Column 4 uses TS1990 commuting zones and Column 5 uses the FKV1990 commuting zones. Standard errors are in parentheses. All coefficients are significant with p-values less than 0.01.

Figure 7: Differences in Effect Based on Cluster Cutoff



(a) Effect by Cutoff Height

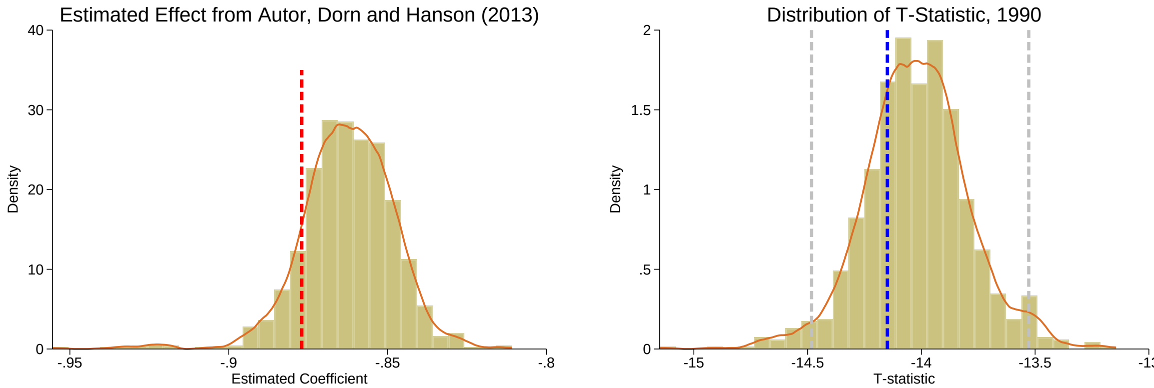
(b) Effect by Cutoff Height, Scaled by IQR

Note: Author's calculations based on replication of Tolbert and Sizer's method. Panel (a) shows estimates of β_1 from Equation 6 for different definitions of commuting zones based on height cutoff, while Panel (b) shows estimates of β_1 scaled by the difference in exposure between the 25th and 75th percentile commuting zone. The horizontal line in Panel (a) is the main estimate from Autor, Dorn and Hanson (2013)

Figure 7 displays the results of this exercise. Panel (a) shows the raw coefficient, and Panel (b) shows the coefficient scaled by the interquartile range of ΔIPW_{uit}, which changes depending on the cutoff. In Panel (a), the red dashed line is the estimate from Autor, Dorn and Hanson (2013).

Our results show that there is some variation in the estimate based on the cutoff value. As with Figure 5, estimates using more clusters have lower standard errors, but are also lower magnitude, possibly due to attenuation bias. Around the cutoff value that most closely replicates TS1990, the estimate is the most negative. However, cutoff values marginally higher or lower give different results, which reinforces the point we made with Figure 2 - the number of clusters that merge is incredibly dense near the cutoff for commuting zones, so that any change in the cutoff changes the commuting zone definitions non-negligibly. This density causes estimates using commuting zone observations to change near the cutoff. Given the sensitivity of estimates to the chosen cutoff, best practices would be to report estimates for a broad range of cutoffs. In this particular example, the estimates are not statistically distinguishable, but this may not be the case in all applications.

Figure 8: Distribution of Effect, 1990-2000



(a) Coefficient (b) T-Statistic

Note: Histogram in Panel (a) plots estimates of β_1 from Equation 6, based on commuting zone realizations as outlined in Section 2. Histogram in Panel (b) plots t-statistics derived from estimating Equation 6, based on commuting zone realizations from Foote, Kutzbach and Vilhuber (2020). Blue vertical dashed line is t-statistic using FKV1990, and gray vertical lines are the 2.5th and 97.5th percentiles of the t-statistic distribution.

4.2.3 Sensitivity to Errors in Flows Data

Next, we show how the estimates are affected by variation induced by resampling the commuting flows based on the implicit distribution from the margins of error in flows. We use the same resampling outcomes as Section 3.2.

We use these different commuting zone definitions to aggregate counties and then estimate Equation 6. These estimates are graphed in Figure 8, which shows the distribution of the estimated effect for the 1990-2000 period. The red vertical dashed line in Panel (a) shows the estimate under the standard commuting zone definitions. The estimates are somewhat dispersed, and the standard deviation is 0.015, which implies that the standard errors are understated by about 20% (comparing to standard error in column 5 of Table 3). Additionally, the distribution is not centered on the red line, which implies that the conventional estimate of the effect overstates the true effect.

Another way to display these results is to show the resulting distribution in t-statistics, which incorporates information on both the coefficient and standard error; we show this distribution in Panel (b) of Figure 8. The blue vertical dashed line is our originally estimated t-statistic, and the two gray vertical lines are the 2.5th and 97.5th percentile of the estimated t-statistics.

While the estimates from Autor, Dorn and Hanson (2013) remain significant when using all the realizations of commuting zones, this exercise demonstrates that there is additional uncertainty induced by the construction of commuting zones that is not addressed in empirical estimates that use these definitions, which may overstate the precision of results.

5 Solutions and Conclusion

Numerous influential papers in labor economics have used commuting zones as an alternative definition to local labor markets. However, researchers typically do not evaluate how the methodology used to construct commuting zones may impact their findings, nor have there been any evaluations of the sensitivity of commuting zones to design feature more generally. Our paper contributes to this literature by analyzing this methodology and its implications for empirical applications.

We document that the commuting zone design methodology is sensitive to uncertainty in the input data and parameter choices and we demonstrate how these features affect the resulting labor market definitions. Furthermore, we demonstrate that the underlying uncertainty of the commuting zone definitions affects empirical estimates that use them as a unit of analysis.

We propose two ways to convey this additional uncertainty. First, to capture the subjectivity of the cutoff choice, researchers should display results for a variety of different cluster counts resulting from a range of cutoff values. This point is particularly important for researchers applying the methodology from (Tolbert and Sizer, 1996) to new datasets or for characterizing labor markets outside the United States, given that cluster counts are subjective and that results can differ considerably based on the chosen cutoff. Second, to capture the underlying uncertainty in the commuting flow data used as an input, researchers should re-estimate results using multiple realizations of commuting zones, and thereby incorporate uncertainty due to the underlying error in the measurement of flows. To assist researchers in implementing these robustness checks, Foote, Kutzbach and Vilhuber (2020) provide the code underlying the analyses reported in this article, which can easily be adapted to other contexts. If using U.S. commuting zones, researchers can download two datasets that provide commuting zone realizations (for the cutoff and resampling methods) and follow the steps outlined in the README file posted with the code.[23]

Our findings are relevant for empirical labor economics as well as other fields using commuting zones and could be extended to area definitions based on other flows data (e.g. trade, commodity, or finance flows). Researchers can validate results by examining either the distribution of parameter estimates or the distribution of t-statistics for the estimates, as described in the previous section.[24]

[23] In addition to Foote, Kutzbach and Vilhuber (2020), the most recent code can be found at github.com/larsvilhuber/MobZ/.

[24] Researchers might also use an alternative flows dataset to define local labor markets. For example, hiring outcomes data give the flow of workers acceding into new jobs by geographic origin and destination. For more information, see the Job-to-Job flows data produced by the Census Bureau's Longitudinal Employer-Household Dynamics program at `https://lehd.ces.census.gov/data/`.

References

Afonso, António, and Ana Venâncio. 2016. "The relevance of commuting zones for regional spending efficiency." *Applied Economics*, 48(10): 865–877. https://doi.org/10.1080/00036846.2015.1088145.

Amior, Michael, and Alan Manning. 2018. "The Persistence of Local Joblessness." *American Economic Review*, 108(7): 1942–70. https://doi.org/10.1257/aer.20160575.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–68. https://doi.org/10.1257/aer.103.6.2121.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies? Books from Upjohn Press*, W.E. Upjohn Institute for Employment Research. https://ideas.repec.org/b/upj/ubooks/wbsle.html.

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, 23(1): 1–76. https://www.brookings.edu/bpea-articles/regional-evolutions/.

Bound, John, and Harry J Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54. https://doi.org/10.1086/209949.

Briant, A., P.-P. Combes, and M. Lafourcade. 2010. "Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations?" *Journal of Urban Economics*, 67(3): 287 – 302. https://doi.org/10.1016/j.jue.2009.09.014.

Bureau of Economic Analysis. 2019. "Table 30: Economic Profile by County, 1969-2018." U.S. Department of Commerce [producer] [Datafile]. https://apps.bea.gov/regional/zip/CAINC30.zip.

Bureau of Labor Statistics. 2020. "Quarterly Census of Employment and Wages – Data Files." https://www.bls.gov/cew/downloadable-data-files.htm (accessed 2020-09-20).

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64. https://doi.org/10.1086/209979.

Casado-Díaz, José Manuel, and Mike Coombes. 2011. "The Delineation of 21st Century Local Labour Market Areas: A Critical Review and a Research Agenda." *Boletín de la Asociación Española de Geografía*, 57: 7–32. https://bage.age-geografia.es/ojs/index.php/bage/article/view/1390.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United

States." *Quarterly Journal of Economics*, 129(4): 1553–1623. https://doi.org/10.1093/qje/qju022.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524. https://doi.org/10.1257/aer.20131706.

Economic Research Service. 2012. "1980 and 1990 Commuting Zones and Labor Market Areas." United States Department of Agriculture [Dataset]. https://www.ers.usda.gov/webdocs/DataFiles/48457/czlma903.xls?v=7728.8 (accessed 2020-09-01).

Foote, Andrew, Mark Kutzbach, and Lars Vilhuber. 2020. "Replication code and data for: Recalculating ... How Uncertainty in Local Labor Market Definitions Affects Empirical Findings ." Labor Dynamics Institute [publisher], Zenodo [distributor], https://doi.org/10.5281/zenodo.4072428.

Foote, Andrew, Michel Grosz, and Ann Stevens. 2019. "Locate Your Nearest Exit: Mass Layoffs and Local Labor Market Response." *ILR Review*, 72(1): 101–126. https://doi.org/10.1177/0019793917753095.

Fotheringham, A.S., and D.W.S. Wong. 1991. "The modifiable areal unit problem in multivariate statistical analysis." *Environment and Planning A*, 23: 1025 – 1044. https://doi.org/10.1068/a231025.

Fowler, Christopher S., Danielle C. Rhubart, and Leif Jensen. 2016. "Reassessing and Revising Commuting Zones for 2010: History, Assessment, and Updates for U.S. 'Labor-Sheds' 1990–2010." *Population Research and Policy Review*, 35: 263–286. https://doi.org/10.1007/s11113-016-9386-0.

Franconi, Luisa, Daniela Ichim, and Michele D'Aló. 2017. "Labour Market Areas for territorial policies: Tools for a European approach." *Statistical Journal of the IAOS*, 33(3): 585–591. https://doi.org/10.3233/SJI-160343.

Halkidi, Maria, Yannis Batistakis, and Michalis Vazirgiannis. 2001. "On Clustering Validation Techniques." *Journal of Intelligent Information Systems*, 17(2): 107–145. https://doi.org/10.1023/A:1012801612483.

Head, Allen, and Huw Lloyd-Ellis. 2012. "Housing Liquidity, Mobility, and the Labour Market." *Review of Economic Studies*, 79(4): 1559–1589. https://doi.org/10.1093/restud/rds004.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251. https://doi.org/10.3982/ECTA4657.

Minnesota Population Center. 2016. "National Historical Geographic Information System." https://doi.org/10.18128/D050.V11.0.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review*, 108(12): 3855–90. https://doi.org/10.1257/aer.20151507.

Moretti, Enrico. 2011. "Local Labor Markets." , ed. O. Ashenfelter and D. Card Vol. 4 of *Handbook of Labor Economics*, Chapter 14, 1237–1313. Elsevier. https://doi.org/10.1016/S0169-7218(11)02412-9.

Niedzielskia, Michael A., Mark W. Horner, and Ningchuan Xiao. 2013. "Analyzing scale independence in jobs-housing and commute efficiency metrics." *Transportation Research Part A: Policy and Practice*, 58(10): 129–143. https://doi.org/https://doi.org/10.1016/j.tra.2013.10.018.

Notowidigdo, Matthew J. 2020. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics*, 38(3): 687–725. https://doi.org/10.1086/706048.

Office of Management and Budget. 2000. "Standards for Delineating Metropolitan and Statistical Areas." Federal Register Notices 65 FR 82227. https://www.federalregister.gov/documents/2000/12/27/00-32997/standards-for-defining-metropolitan-and-micropolitan-statistical-areas.

Office of Management and Budget. 2010. "2010 Standards for Delineating Metropolitan and Statistical Areas." Federal Register Notices 75 FR 37245. https://www.federalregister.gov/documents/2010/06/28/2010-15605/2010-standards-for-delineating-metropolitan-and-micropolitan-statistical-areas.

Restrepo, Pascual. 2015. "Skill Mismatch and Structural Unemployment." Massachusetts Institute of Technology Job Market Paper. https://economics.mit.edu/files/11078 (accessed 2020-09-20).

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*, 90(6): 1257–78. https://doi.org/10.1086/261120.

Schmutz, Benoit, and Modibo Sidibé. 2014. "Job Search and Migration in a System of Cities." Centre de Recherche en Economie et Statistique Working Papers 2014-43. https://ideas.repec.org/p/crs/wpaper/2014-43.html.

Tolbert, Charles M, and Molly Sizer. 1996. "US commuting zones and labor market areas: A 1990 update." Economic Research Service.

Tolbert, Charles M, and Molly Sizer Killian. 1987. "Labor market areas for the United States." *ERS staff report.*

U.S. Census Bureau. 1992*a*. "1990 Census of Population and Housing Public Use Microdata Samples (PUMS): United States." Department of Commerce Technical Documentation D1-D90-PUMS-14-TECH. https://www2.census.gov/prod2/decennial/documents/D1-D90-PUMS-14-TECH-01.pdf (accessed 2020-09-20).

U.S. Census Bureau. 1992*b*. "Census of Population 1990: Number of Workers by County of Residence by County of Work." Department of Commerce Summary Tape File S-5.

U.S. Census Bureau. 2014. "American Community Survey: Multiyear Accuracy of the Data." Department of Commerce Technical Documentation. https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/MultiyearACSAccuracyofData2013.pdf (accessed 2020-09-20).

U.S. Census Bureau. 2017*a*. "1990 County-to-County Worker Flow Files." U.S. Department of Commerce [producer] [Datafile]. https://doi.org/10.3886/E100617V1 (accessed 2020-09-07).

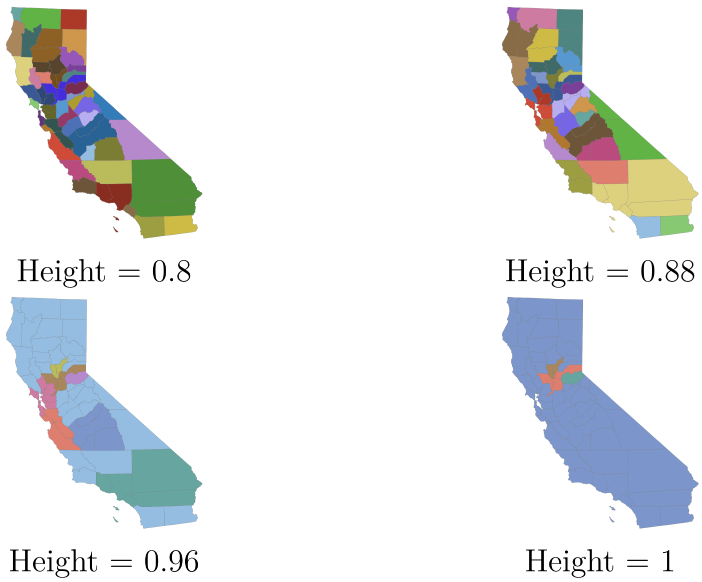
U.S. Census Bureau. 2017*b*. "2009-2013 5-Year American Community Survey: Commuting Flows." https://doi.org/10.3886/E100616V1 (accessed 2020-09-08).

Wozniak, Abigail. 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*, 45(4): 944–970. https://doi.org/10.3368/jhr.45.4.944.

Yagan, Danny. 2016. "The Enduring Employment Impact of Your Great Recession Location." University of California, Berkeley Working Paper.

Tables and Figures Appendix

Figure 9: Clusters in California at Incremental Height Cutoffs



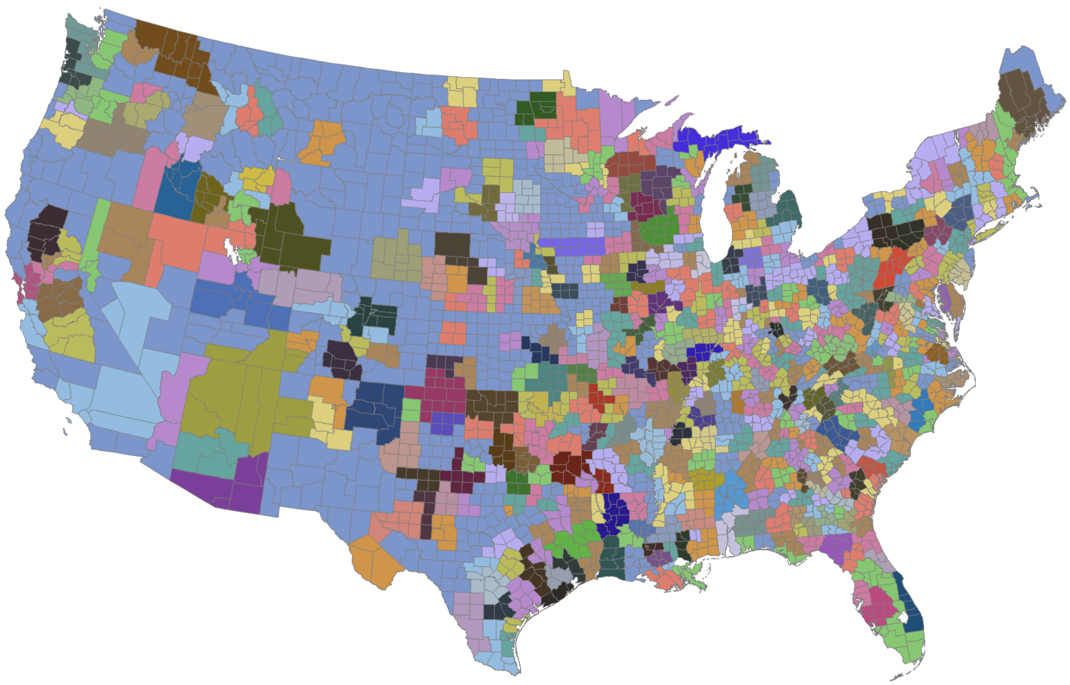
Notes: The above graphs are generated using the methodology outlined in Section 2, using 1990 Census Journey to Work data. More detail is in the text in Section 2.3.

Table 4: Summary Statistics of Ratio of MOE to Flows

	Mean	25th Pctile	50th Pctile	75th Pctile
All counties	1.236	0.845	1.370	1.600
Flows <100	1.432	1.148	1.500	1.636
Flows 100-1000	0.444	0.301	0.414	0.549
Flows 1000-10000	0.131	0.087	0.124	0.169
Flows 10000+	0.037	0.024	0.036	0.049

Notes: Author's calculation using Margins of Error and total flows from 2009-2013 5-year ACS Journey to Work data. See Section 3.2 for usage details.

Figure 10: Hierarchical Clustering, Cutoff = 0.945



Notes: The above cluster mapping was generated using the methodology outlined in Section 2, using 1990 Census Journey to Work data, with a height cutoff of 0.945, which is above FKV1990. The dispersed set of light blue counties are all in the same cluster, even though parts of the cluster are non-contiguous; there are 292 counties in that cluster.

Table 5: Summary statistics for empirical example

	TS 1990 CZs	Replicated FKV1990 CZs
Log UI receipts	9.23	9.59
	(1.89)	(1.54)
$Bartik_{it}$	0.006	0.006
	(0.022)	(0.022)

Notes: For details on construction and source, see Section 4. Standard deviations in parentheses.

ONLINE APPENDIX (not for publication)

Full-size maps

Figures 11 and 12 show full size versions of the maps of commuting zones and replicated commuting zones reported in the paper.

Figure 11: Replication of Commuting Zones, FKV1990

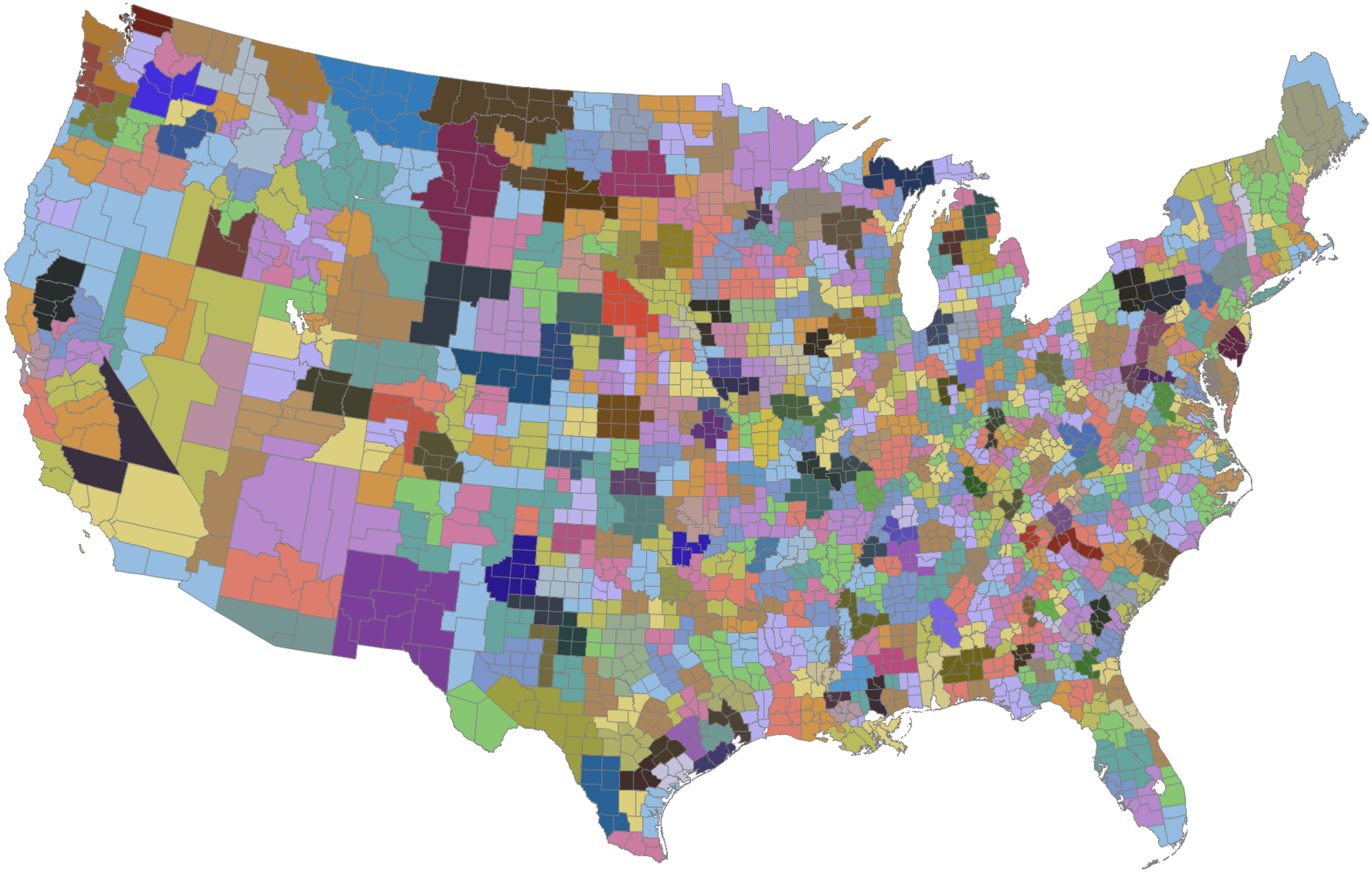
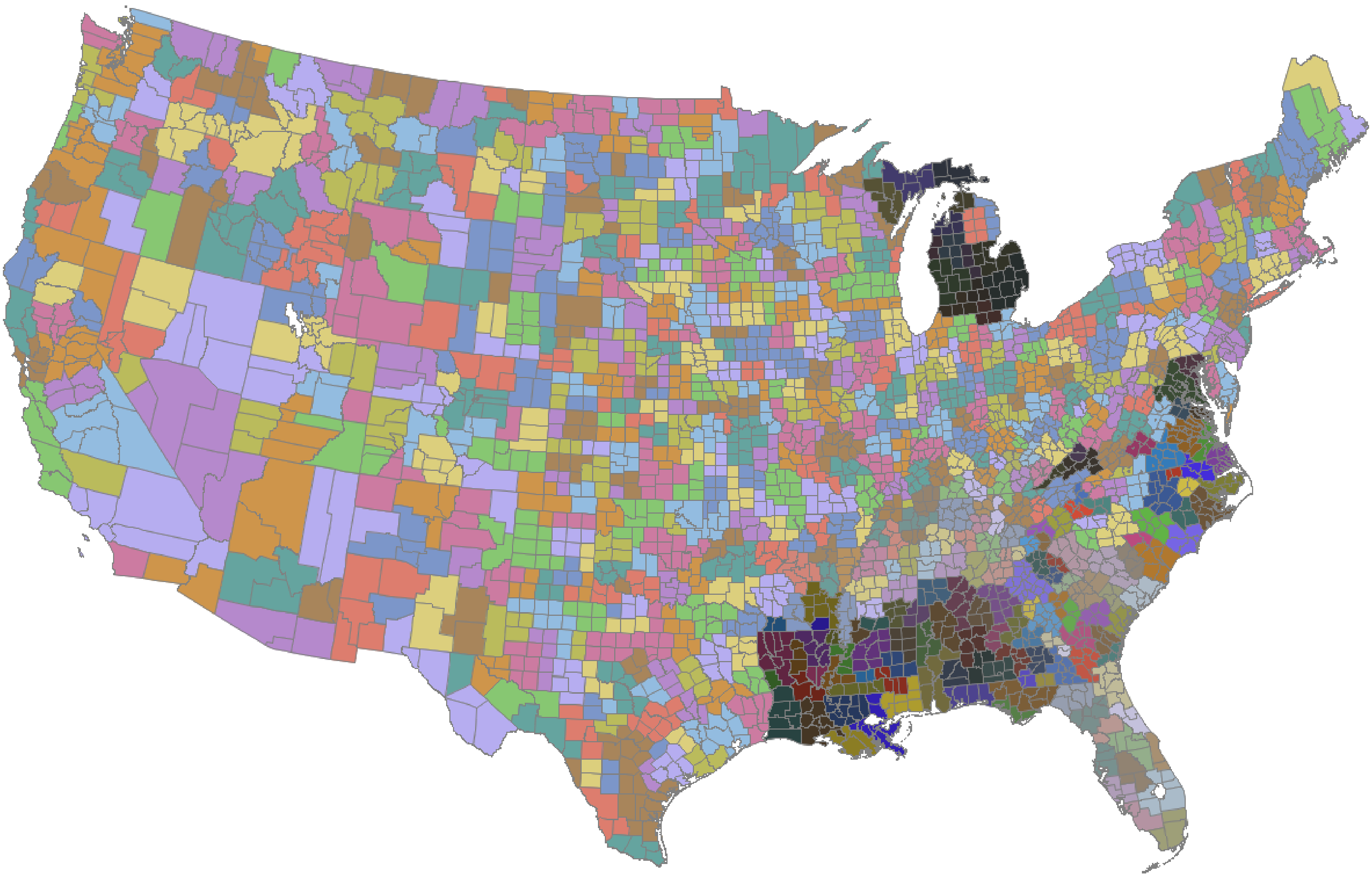


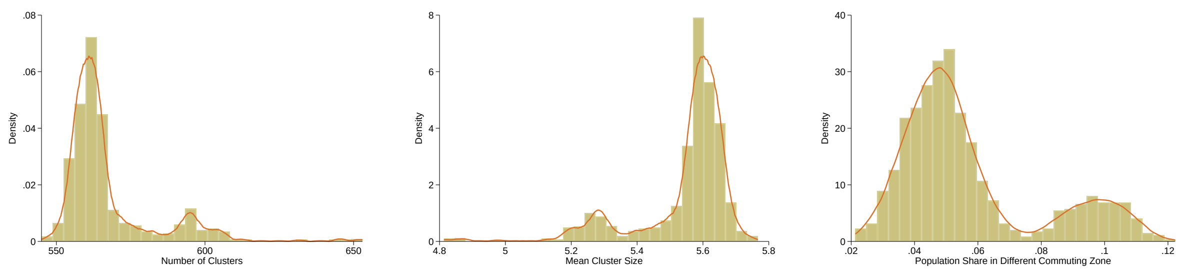
Figure 12: Commuting Zones from Tolbert and Sizer (1996)



Robustness check for sampling procedure

We replicate the analysis reported in Figure 4 using the 2009-2013 ACS data. This has the additional advantage of using the MOEs specifically constructed for the data, rather than projecting them onto the 1990 commuting flows. The results are shown in Figure 13.

Figure 13: Results from Re-sampling Commuting Flows, JTW 2009-2013



(a) Number of clusters

(b) Average Number of Counties

(c) Share of Mismatched Population

Notes: Results from the resampling procedure described in Section 3.2, using the JTW data from 2009-2013.