Popul Res Policy Rev (2016) 35:263–286
DOI 10.1007/s11113-016-9386-0

Reassessing and Revising Commuting Zones for 2010: History, Assessment, and Updates for U.S. 'Labor-Sheds' 1990–2010

Christopher S. Fowler[1] · **Danielle C. Rhubart**[2] · **Leif Jensen**[2]

Received: 12 August 2015 / Accepted: 26 February 2016 / Published online: 7 March 2016

Abstract This paper employs commuter flow data from the 1990 and 2000 Decennial Censuses, and the 2006–2010 American Community Survey to replicate, evaluate, and extend the delineation of commuting zones first proposed by Tolbert and Killian (Labor Market Areas for the United States, 1987). Commuting zones offer a valuable tool for research on regional economies and have long served rural sociologists, economists, and geographers interested in a representation of the economy that acknowledges a connection between urban and rural areas and the capacity of economic systems to cross state lines. Our delineations provide both an update in the form of new delineations for 2010 and a revised set of 1990 and 2000 delineations that benefit from a consistent methodology across decades. We also provide a set of tools for comparing delineations across methods and over time. In presenting our revised delineations, we shed light on the role of expert opinion in the original delineations, the strengths and weaknesses of the original method, and offer suggestions for further revision of this tool that may better reflect the theoretical conception of commuting zones.

Keywords Commuting zones · Regional economies · Clustering · Geographic boundaries

✉ Christopher S. Fowler
csfowler@psu.edu

[1] Department of Geography, Population Research Institute, The Pennsylvania State University, 302 Walker Building, University Park, PA 16802, USA

[2] Department of Agricultural Economics, Sociology, & Education, Population Research Institute, The Pennsylvania State University, Armsby Building, University Park, PA 16802, USA

Introduction

In the 1980s, researchers at the USDA's Economic Research Service (ERS) and at land grant universities from around the country undertook a task to overcome problems associated with delineating geographic areas of economic integration. Then—as now—the county-level delineations were inadequate for measuring areas of integrated labor processes. Counties are the most common geographic unit for the provision of economic data but they do not begin to capture the geographic extent of regional labor markets that extend well beyond county and often state borders. Various census-defined urban categories (e.g., consolidated metropolitan statistical areas, metropolitan statistical areas) might represent an alternative; however, they change in geographic scope between censuses and exclude rural counties that may, nevertheless, be closely tied to urban labor markets, and are not widely used for reporting economic data. From the perspective of researchers interested in rural areas, the exclusion of broad geographic areas that may not have had an urban core was also problematic. Faced with this set of complex limitations, ERS and its partners worked together to develop 'Commuting Zones' (CZ's) and 'Labor Market Areas'(LMA's) (Tolbert and Killian 1987). These geographic delineations, based initially on 1980 Census long-form data about county of residence and county of work, filled several important gaps:

1. based on counties they could be easily combined with numerous other data sources for which county data were available,
2. county geographic definitions change more slowly than most other administrative boundaries so the units of analysis would remain relatively consistent across years,
3. covering the entire U.S., they integrated rural and urban populations into a consistently defined set of observations, and
4. covering all counties in the U.S., they offered a continuous spatial surface (at least in the contiguous 48 states) suitable for a range of spatial analytic techniques.

The delineations proposed by Tolbert and Killian were successful and have been used for research in academia and for program and public policy formulation (Economic Research Service 2015). The success warranted a second delineation after the 1990 Census, and a revision of the CZs after the 2000 Census. CZs continue to play a role in research into the present day (Autor and Dorn 2009; Fowler and Kleit 2013; Martin et al. 2015; Molloy et al. 2011). As the 1990 and 2000 delineations have grown dated, CZs have received less utilization (Economic Research Service 2015). The LMAs, primarily meant as a tool for accessing special Census microdata tabulations for 1980 and 1990, are now completely out of date. In contrast, the problems that these delineations were meant to overcome (representations of regional economies and continuous spatial surfaces for analysis in particular) have received increased attention in the literature (Feser and Isserman 2005; Feser 2003; Fowler and Kleit 2013; Lloyd et al. 2014; Voss 2007). This paper is meant to serve as a starting point for an update of the CZ delineations to facilitate their use in ongoing research. Specifically, this paper provides:

1. A background on the history and methodology behind the original delineations.
2. An evaluation of the strengths and weaknesses of the original methodology including an assessment of the extent to which expert input shaped the original delineations.
3. A proposed update of the delineations using a purely statistical method calibrated to the original delineations.
4. Updated delineations using 1990, 2000 decennial census and 2006–2010 ACS data.
5. Tools for comparing both the original and newer delineations across censuses.
6. Reproducible code for generating and comparing delineations across methods and over time so that researchers can continue to update CZs as new data and new methods for estimating cluster membership become available.

Ultimately, our analysis of the original methodology indicates a need for a substantial revision of CZs to include more robust measures of fit both for CZs and for the counties that belong to them. While the delineations we provide here will be suitable for researchers interested in comparing results against earlier findings and as an established baseline for further development of CZ delineations, we conclude by arguing for a revision of the CZ methodology.

The paper continues in four parts. First, we offer some background on the development of CZs and LMAs and some of the key applications of this tool in rural sociology, geography, regional science, and demography. Next, we turn to a methodological discussion of the original delineations including an assessment of the strengths and weaknesses of that delineation and the application of expert information to the final outcome. Subsequently, we offer information about the methods we employ for extending the delineation to 2006–2010 ACS data and, in particular, calibrating our own methods to the original. In concluding, we argue that the delineations presented here are important for purposes of comparability and continuity with respect to prior and ongoing research efforts. Nevertheless, the underlying methodology for defining CZs is flawed and in need of revision. To this end we offer some suggestions for improvement and describe ongoing efforts to revise the methodology. The work presented here both supports the use of CZs in the context of research requiring continuity and lays the groundwork for understanding how and why the CZs require revision.

Given its intended role as a clearinghouse for informed use of the updated CZ delineations, this paper conforms to best practices in reproducible research (Gentleman and Lang 2007), and all of the input data and code necessary for the replications presented here as well as spreadsheet and GIS compatible files with the old and updated delineations are available from the corresponding author on request.

How Commuting Zones and Labor Market Areas Came to Be

As a recognized spatial delineation, CZs and LMAs emerged in the late 1980s as a result of a collaborative project between the Rural Labor Market Section of ERS and researchers at land grant universities that were part of Regional Research

Project S-184, "Labor Markets and Labor Differentiation in Nonmetropolitan America." The effort to delineate CZs and LMAs benefitted tremendously from the input of Calvin Beale. A geographer at ERS, Beale was singled out as indispensable "for sharing his impressive knowledge of this country and his extensive evaluation of these labor market areas" (Tolbert and Killian 1987, p. iii). It is Beale who primarily provided the expert advice on ad hoc decisions regarding the assignment of counties mentioned elsewhere in this paper.

This project served two purposes: it delineated CZs that were to represent 'commute sheds' by grouping counties together based on commuting patterns through a clustering algorithm, and it then grouped CZs together into LMAs such that each LMA had a minimum of 100,000 residents. For most parts of the country, CZs and LMAs were one and the same. The exception was in geographies with small population sizes, where CZs did not meet the minimum population threshold and so had to be combined. The CZs that were combined to meet that threshold and the CZs that already met the threshold, together, comprised the LMAs. The first task was intended as an improvement on the geographic units available for studying regional economies, while the LMA's were undertaken in support of a special Census tabulation of microdata from the 1980 U.S. Census of Population and Housing, the so-called PUMS-D.[1] This 1 % sample of the U.S. population included social, demographic, and economic indicators for households and all individuals residing within them. By including LMA as a household-level variable, researchers were able to document and explore rural and urban labor market characteristics and their implications for household circumstances (e.g., poverty) and individual outcomes (e.g., employment). In a follow-on regional research project (S-229) titled, "The Changing Structure of Local Labor Markets in Nonmetropolitan Areas," many of the same researchers continued with analysis of the 1980 PUMS-D data, but also replicated their methodology with 1990 Census data and worked to create the PUMS-L as a companion to the PUMS-D (Tolbert and Sizer 1990).

The rationale for this effort was recognition of the inadequacy of counties as units to understand an area's economy. As stated by Tolbert and Killian, "[a] local economy and its labor force are not bounded by the nearest county line, but by interrelationships between buyers and sellers of labor" (Tolbert and Killian 1987, p. 1). As many on the team were rural social scientists, they were motivated by the inadequacy of nonmetropolitan areas for capturing the diversity of rural America and by the need for a geographic standard to capture labor markets. As noted, ultimately the goal was to provide a better understanding of the implications of labor market context for individual circumstances and outcomes.

Our work here focuses on extending Tolbert and Killian's delineation of CZs through 2010 to support consistent analysis of regional economies over time, but we do not seek to replicate LMAs. CZs represent the conceptual target of the original research program, while LMAs were more of a convenience measure driven by

[1] Commuting zones provided units of analysis that represented regional economies. However, a number of rural CZs failed to meet the 100,000 minimum population threshold required for use in a microdata tabulation. These CZs were aggregated to meet the population threshold. The original CZs along with the aggregated CZs are called LMAs. While PUMS files using LMAs would be comprehensive of all geographic areas of the US, they mask nuanced delineations of rural economies.

Census confidentiality requirements. ERS replicated its CZ delineations in 2000, but did not seek to update LMAs. The Census also dropped the long-form survey and associated microdata products in favor of the American Community Survey (ACS) after 2000 so a repeat of the PUMS-style microdata products is not a simple matter to pursue. With these caveats in mind, we do think there is value in a Census microdata product that offers rural scholars the same opportunities for analyzing microdata as those afforded by the PUMS-D and PUMS-L products, and future research will explore the potential to leverage ACS data and Census restricted data to achieve something similar.

Previous Uses of Commuting Zones and Labor Market Areas

Since their first creation, CZs and LMA delineations have proved to be valuable to both government and academia. Killian and Tolbert (1993) emphasized the role of CZs and LMAs in better understanding the implications of context on individual outcomes. They stated:

> One major reason for incorporating space into our conceptual models of social and economic inequality is the assumption that the structural context of the local labor market will have a unique, independent impact on individual social and economic processes. … Therefore, in identifying local labor market areas, we want to construct a geography that can be used with large data files that permit detail on individuals and their places of residence and places of work. (Killian and Tolbert 1993, p. 75)

While this paper does not provide a comprehensive overview of the academic pieces that have benefited from the use of CZs and LMAs, a sense of the range of use can be gained by mentioning a sampling. Much of this work has used CZs and LMAs as a measure of 'labor sheds'; economically connected areas where individuals can live and work and where, like the watersheds after which they are named, few residents are assumed to cross between sheds for employment purposes. The need to understand the implications of labor shed characteristics for job growth has been long recognized. Research by Killian and Parker (1993) used CZs to show that the effect of education on local employment growth had been previously overstated by the literature. In fact, they show that local industries, labor costs, labor force skills, and location are also important factors in determining employment growth, such that education is only a small component for stimulating employment growth. Research by Tickamyer and Bokemeier (1987) using the LMAs available in the PUMS-D of the 1980 Census revealed how the industrial composition of labor sheds affects men's and women's earnings in different ways, with penalties for women most apparent in LMAs dominated by mining vis-à-vis agriculture. That this work was focused on Kentucky and its environs underscores how labor shed characteristics can have significant effects even within circumscribed regions. Later work by Davis and colleagues (2003) used CZs to measure the impact of local job growth on the probability of poor jobless adults finding a job in Oregon. The authors show that local job growth increases the likelihood of poor jobless adults getting a job (Davis et al. 2003). Allegretto and colleagues (2009) used CZs for measuring the spatial heterogeneity of teen employment estimates. They find that past research has failed to account for the spatial heterogeneity of employment patterns and have, as a result, overestimated the effect of state minimum wage policies on teen employment (Allegretto et al. 2009). And most recently, Dupor and McCrory (2014) have generated CZs to assess labor market changes and their association with spending under the 2009 American Recover and Reinvestment Act. Their findings point to spillover effects in labor market areas during the first 2 years of the Act's implementation. These studies are examples of how CZs and LMAs played a pivotal role in explaining the impacts of labor shed characteristics.

These units of measurement have also accelerated our understanding of multilevel relationships between individuals, households, and labor markets. Cotter (2002) made use of labor market delineations in multilevel models to show that both the compositional and contextual factors contributed to the variations in poverty rates between metro and nonmetro areas. Other work analyzing the PUMS-L has stressed that individual characteristics are more important than labor shed variables in shaping unemployment and labor market outcomes generally (Zolnik 2010). That much of this work is of an older vintage reflects the failure to produce a special PUMS from the Census beyond 1990. In fact, more recent work by Fullerton and Villemez (2011) still reaches back to the 1990 PUMS-L file, while Dorn (2009) sought to overcome this limitation by developing a technique to probabilistically match microdata to CZs.

While the aging of the LMA files has led to a marked slowing in the use of CZs and LMAs within rural sociology, the utility of CZs for regional economic analysis and the burgeoning literature incorporating spatial processes has seen a continued use of CZs in regional science, geography, and sociology. For example, Gibbs and Bernat (1980) used CZs to examine how earnings vary between workers in and outside of industry clusters. Among rural counties, they found that wages tended to be higher for workers in an industry cluster compared to workers not in an industry cluster. Autor and Dorn (2009) utilized the CZ delineations to show how middle skill employment had changed over previous decades and, in particular, how middle skill intensive labor markets had "hollowed-out" beginning in 1980 through the mid-2000s. And most recently, Autor and Dorn used CZs again to examine growth in low-wage jobs within labor markets that had specialized in tasks that could be easily routinized (D. H. Autor and Dorn 2013). CZs have also been used for measuring migration. While Molloy et al. (2011) used CZs to measure internal migration, Tolbert et al. (2009) used the delineations to model migration during the 1985–1990 and 1995–2000 time periods and found that individual-level predictors of migration between CZs were relatively consistent across the two time periods observed.

Regional economics research has also utilized CZs and LMAs to better understand changes in firm formation and entrepreneurship. Armington and Acs (2010) used LMAs to examine new firm formation and how the birth of new firms is related to industrial regions and technological entrepreneurship. CZs have also been used to understand the local area attributes that are associated with food manufacturing plant investments in the Corn Belt region of the US (Henderson and McNamara (2000). The authors specifically use CZ level measures of wages and unemployment to account for cross-county commuting of workers. CZs have also been used to analyze industry clusters where linkages between firms could reasonably be expected to occur across counties within the same labor market (Fowler and Kleit (2013).

While much of this review has covered academic uses of CZs and LMAs, policy literature has also relied on these delineations. A recent prominent example came from a series by the Urban Institute based on an online mapping tool, (Martin et al. 2015; Nichols et al. 2015; Pendall et al. 2015), which employs CZs to examine a wide range of demographic characteristics. In the justification for their use of CZs, the authors say

> Though the counties within a commuting zone might not follow identical trends in birth, mortality, or migration rates, the commuting zone's demographic trajectory can plausibly be projected as the joint development of a single local economy (see Tolbert and Sizer 1996). Hence, use of the commuting zone as a unit of analysis makes sense for population projections. Further, the projections generated for these commuting zones can then be applied to the discussion of economic outcomes (such as housing markets) as well as social or other outcomes, where the commuting zone also serves as a natural unit of social analysis (Hildner et al. 2015, p. 3).

Commuting Zones for 2010

Looking forward, CZs have the potential to see increased use for their capacity to capture regional economies while offering a contiguous spatial surface. An upsurge in the use of spatial models for sociological, economic, and demographic research means that an increasingly large body of work depends on the development of spatial weights matrices that require contiguity of observations. Researchers who might otherwise use metropolitan areas or other similar Census designations for their units of observation will find that key relationships between metropolitan areas, so crucial in coastal areas in particular, are easily specified when using CZs, and laborious in the context of numerous metropolitan area 'islands.' Moreover, as the spatial structure of employment changes, it is increasingly reasonable to group surrounding rural areas with sometimes distant urban cores.

The concept of a CZ has significant current relevance and the delineations constructed since 1980 have sustained a broad set of research projects. This history justifies extension of the delineations to 2010 to allow researchers to take advantage of current data and to reflect changes in connectivity between counties in the past three decades. Nevertheless, the original delineation was undertaken in an era with substantially different computing resources and with the benefit of expert input shaping the ultimate designation of counties within CZs. To our knowledge, no study has sought to examine the role of expert input in the original delineations or considered the degree of change in the delineations over time. In the following sections, we describe mechanisms for replicating and extending the CZ delineations as well as some descriptive mechanisms for examining change over time and the role of expert opinion.

Data and Methods

The analysis presented here (as with those conducted previously by Tolbert and colleagues) was undertaken using county-to-county commuter flow information compiled by the Census. Initially, this information was developed out of the Census long-form questionnaire (question 22 in both 1990 and 2000 surveys) that asks "at what location did this person work LAST WEEK?" (United States Census Bureau 2010b emphasis in the original). More recently, this question has become part of the American Community Survey (ACS). In all cases, the respondent is asked to provide the location of the place where they worked the most hours last week including the address, county, and state. The Census then geocodes this response to the 'place' level and ultimately to the block level (McKenzie 2013).

The ACS represents a significant departure from the Census long form of previous decades and merits some additional consideration here. The ACS is collected from a smaller sample of individuals and the survey is administered continuously so that estimates are developed as summaries over multiple years instead of single point estimates as in the 1990 and 2000 long-form responses (We use the 2006–2010 5-year summary data here). The ACS estimates also differ in that they include a margin of error along with the flow estimates. In many counties with smaller populations, this margin of error can be significant relative to the size of commuting flows and may include zero. Of the 136,794 county commuting pairs identified in the ACS data for 2006–2010, 96,314 or 70 % have a margin of error that includes zero. The margin of error reflects the smaller sample size of the ACS as well as efforts to protect the privacy of individuals surveyed for the ACS. More optimistically, the total commuting flow estimates from county-to-county pairs where the margin of error includes zero represent just a little over 1 % of the total commuting flows captured in the ACS data, so its impact on the results, while real, should not drive outcomes even as it is important for some small counties. Only sixty-two of 3143 counties have ten percent or more of their flows in these problematic count areas and only one county, Loving, Texas has more than twenty-five percent of its flows in these count areas. Loving, with its 2010 population of eighty-three (U.S. Census 2010a) is likely going to be a problem with any data source seeking to protect privacy.

For this analysis, we use the point estimates provided by the ACS without accounting for the margin of error. This is a limitation of this study that we hope to overcome in future work where we will base our analysis off of the restricted use ACS file where the actual county-to-county flows will be available as well as additional information about the uncertainty around the point estimates. In the present case, we might attempt to incorporate the uncertainty represented by the margin of error as a form of robustness check, but it is not clear how much additional information this might grant us given the limitations, discussed below, in judging the quality of commuting zones.[2]

[2] We also conducted a robustness check using the commuting flow data associated with the longitudinal employer-household dynamics (LEHD) data (Abowd et al. 2009). LEHD Origin–Destination Employment Statistics (LODES) data have apparent advantages over the ACS in that it is based on a larger population than the ACS sample and does not include a margin of error. However, as discussed by Spear

In addition to the Census commuter flow data, this analysis also relies on geographic boundary definitions for 1990, 2000, and 2010 and the original ERS CZ delineations. The Census geographic information was made available through the National Historical Geographic Information System (Minnesota Population Center 2011). ERS makes the original CZs available for download on their web site (Economic Research Service 2015).

After removing records for Puerto Rico and other outlying U.S. territories, the 1990 and 2000 files from NHGIS and ERS contain 3141 county records, while the 2010 NHGIS data contain 3143.[3] Comparing data files across sources using federal information processing standards (FIPS) codes for counties reveals that these sources are comparable within years, but contain some well-known changes across years. First, the boroughs in Alaska exhibit significant changes across all three censuses. In total, seven boroughs are reconfigured or renamed across the three data points. In Florida, Dade County becomes Miami-Dade County between 1990 and 2000 with a change in FIPS code to preserve alphabetical order. Two Virginia counties disappear with South Boston merging into surrounding Halifax County between 1990 and 2000 and Clifton Forge merging into surrounding Allegheny County between 2000 and 2006. Yellowstone National Park is merged into Gallatin and Park counties between the 1990 and 2000 censuses. Finally, Broomfield County, Colorado is created primarily out of Boulder County, but also parts of Adams, Jefferson, and Weld Counties between the 2000 and 2006–2010 observations.

For the purposes of comparability across years in this analysis, we exclude Alaska and merge Broomfield, Clifton Forge, and South Boston counties into Boulder, Allegheny, and Halifax Counties, respectively. To be clear, the cluster delineation files associated with this paper retain the county FIPS codes appropriate to specific years, and these merges are only applied in cases where we attempt to compare delineations across data sources. Subsequent users of the delineation files will be able to use the decade-specific county configuration for cross-sectional analysis and the converted data only when seeking to make longitudinal comparisons.

Footnote 2 continued
(2011) and Graham et al. (2014), the LODES data assign flows to origin–destination pairs based on a stochastic model to protect privacy. In Spear's analysis, this produces a very large number of very small flows and a configuration that is markedly different from the ACS or prior Census long-form products. In our own analysis, clustering using LODES data produced markedly different results from all other data sources, grouping counties into only 318 clusters as opposed to 610 for the 2006–2010 ACS data and 641 for the 2000 Decennial data. As a result, we feel that the ACS data, with all their flaws, are a better choice for this application. As with all of the elements of this analysis, the results of this robustness check are available from the corresponding author.

[3] The apparent similarity between the 1990 and 2000 data includes two Virginia counties that disappear between decades and two Alaska boroughs that appear between decades.

Tolbert and Killian's Proportional Flow Methodology

Tolbert and Killian's methodology based on proportional flows between counties was used in both the 1987 delineation[4] (Tolbert and Killian 1987) and the 1996 update[5] (Tolbert and Sizer 1990). The proportional flows measure emphasize reciprocity between counties calculated as follows:

$$\frac{C_{ij} + C_{ji}}{\min\left(W_i, W_j\right)}, \tag{1}$$

where c_{ij} and c_{ji} represent the counts of commuters leaving county i for county j and leaving county j for county i, respectively, and w_i and w_j are the total workforces of those two counties. In some cases (when examining the linkages between a county with a small workforce and a county with a large workforce, in particular), the proportional flow measure will exceed 1 and in this case the measure was uniformly reduced to .999; this indicates the maximum level of connectivity possible between two counties. In preparation for use in statistical clustering techniques, the proportional flow matrix was then further transformed into a dissimilarity matrix by subtracting the proportional flow from 1 and setting the diagonal (the relationship of a county to itself) to zero. In this way, the counties with the strongest links are assigned values of .001 ($1 - .999 = .001$) and counties with no connection are assigned values of 1.

The previous delineations for 1980, 1990, and the internal ERS delineation for 2000 were constrained by software and computing limitations of the day and broke the 3141 county × 3141 county dissimilarity matrix into six overlapping sub-matrices representing different U.S. regions: West, Southwest, Midwest, Central, Northeast, and Southeast.[6] This separation by region had theoretical as well as practical advantages because it allowed the researchers to assess linkages differently in the context of the much more spatially extensive Western counties and the smaller and, often, more densely populated Eastern counties. As described previously, the research team performed their analysis on these sub-groups and then used expert input to allocate counties grouped with different counties in two or more regions to arrive at a national file with each county assigned to a single group.

The actual grouping mechanism employed by Tolbert and colleagues is hierarchical cluster analysis (Kaufman and Rousseeuw 2009), a widely used data reduction technique that groups pairs of 'clusters' based on the strength of the linkage between them. In this application, each county begins as its own cluster, the

[4] Cooperative Agreement S-184: "Labor Markets and Labor Differentiation in Nonmetropolitan America."

[5] Cooperative Agreement S-229: "The Changing Structure of Local Labor Markets in Nonmetropolitan Areas."

[6] There is considerably less documentation on the method employed for the internal ERS delineation conducted in 2004. ERS on its website states that "The identical methodology was used to develop CZs for all three decades" (Economic Research Service 2015); however, software and hardware limitations would not have been a problem in 2004, and personal communication with ERS staff (July 23rd, 2015) indicates that at least the 'expert review' portion of the process was not undertaken for the 2000 delineations. Other differences may also exist.

two counties with the strongest linkage are paired to create a new cluster, then the linkages are recalculated and the two clusters with the next strongest link are paired. The process continues until all the counties are assigned to a single cluster. The art of hierarchical cluster analysis comes in determining where in this process to stop.[7] In the original delineation, Tolbert and colleagues used .98 as the cutoff threshold, a level that led them to arrive at 741 clusters in 1990 and 709 clusters in 2000. There is nothing magical about the number .98, it represents a judgement call that a given level represents an appropriate tradeoff between county fit within clusters (increasing the number of clusters means that on average counties will fit better within their clusters) and clusters that are large enough to actually represent labor sheds. As we will see below, the seemingly straightforward relationship between a .98 cutoff and a delineation of 741 distinct clusters masks a considerable amount of additional input not fully explicated in the papers describing the methodology at the time.

Replicating Tolbert and Killian

A key motivation for this paper is to replicate Tolbert and Killian's work for 1990 and the internal ERS delineation for 2000 as a step towards extending their work into a methodologically consistent definition of CZs for 2010 and beyond. This requires the development of a methodology that operates consistently with the intent of the earlier work but independently of expert input. In this section, we lay out the methodological alternatives we have developed to accomplish this task while also identifying points in the process where expert opinion likely influenced the original delineation. Our presentation here has two parts: the first explains a methodology that includes breaking out counties into the six overlapping regional sub-groups employed in the original methodology, and the second offers a simpler technique that operates on the full set of U.S. counties. We reserve comparison of the different methodologies for the subsequent section of this paper.

Replication with Regional Sub-groups

Tolbert and Killian originally generated CZs within six overlapping regions (Tolbert and Killian 1987 see Table 6 for details). They did this primarily in response to technical limitations of the computer applications of the time (their method kept the number of counties in each cluster analysis below 1000), but they also argue convincingly that the heterogeneity in county size between the Eastern and Western U.S. further justifies this choice. We explore the significance of this decision in the analysis below, but for now it is sufficient to note that the overlapping regions they employ (an additional one state buffer extending around the periphery of each of their sub-regions) are meant to ensure that counties just beyond the edge of a sub-region are not excluded from their appropriate clusters.

[7] Hierarchical cluster analysis is highly sensitive to the cutoff point selected which determines the number of clusters in the analysis. For the purposes of this paper, we are trying to develop a mechanism that allows for cross-decade comparisons against the original delineations.

One issue that arises with the use of overlapping sub-regions is that counties in the overlapping areas may be assigned to one cluster in one sub-region and to a different cluster in another sub-region. Some counties in the center of the country where the sub-regions overlap the most can be assigned to as many as four different clusters. While the original papers reporting on the delineation process provide little detail, we suspect that this was the point where expert advice was most crucial to determining the final delineations. In our analysis of the 1990 data using the same overlapping sub-regions defined as in the original analysis, we initially have only 1764 counties assigned unambiguously to a single cluster. An additional 934 counties are assigned to multiple clusters in different sub-regions, but those clusters are the same in all sub-regions. This means that of 3141 counties 443 counties need to be adjudicated because they are assigned to substantively different clusters across at least two sub-regions.[8]

In replicating the original methodology, we chose to develop a systematic mechanism for adjudicating counties assigned to multiple distinct clusters. Our method iteratively adds the counties with the clearest fit relative to alternatives to the appropriate cluster, recalculating cluster fit statistics after each assignment until all counties are assigned to a cluster. To be clear, our ordering is not based on the county with the best fit for one of the clusters, but the county that shows the best fit *relative to other clusters to which it has been assigned.* Specifically, once all the unambiguous cluster assignments are processed (all but 443 counties in 1990), we calculate the average dissimilarity of the remaining counties from the members of the clusters to which they have been assigned. We compare the average dissimilarities across clusters and the county that shows the biggest difference in dissimilarity between two clusters to which it is assigned is selected and assigned to its best-fitting cluster. After each county is assigned to an unambiguous cluster, the average distances are recalculated for affected counties (to take into consideration that a county has either been removed or added to the clusters they are part of) and the process repeats.

Replication Without Regional Sub-groups

Given the complexity of dealing with ambiguous cluster assignments in the method just described, we also ran our cluster analysis on the complete set of U.S. counties. Hardware and software limitations that drove the choice to use sub-regions in 1987 are no longer relevant, and the unambiguous assignment of counties to clusters from the outset has much to recommend it in terms of methodological simplicity. The only real methodological decision in this replication is what to use as a cutoff point for the hierarchical cluster analysis. With the full set of counties in our dataset, we would expect to use a slightly lower cutoff point to replicate, approximately, the intent of the .98 cutoff used by Tolbert and Killian.[9] We employ a value of .977 in our replication as that value appears to do a reasonable job of replicating the results

[8] Using the same approach 528 counties required adjudication in 2000 and 579 in 2010.

[9] Higher values for the cutoff produce a smaller number of clusters. Lower cutoff values produce a larger number of clusters. To get the same number of clusters in a larger dataset requires a lower cutoff value.

of the original methodology. The mechanism for determining a "reasonable job" is the subject of the next section.

Methods for Comparing Delineations

In order for the delineations covered here to be useful for researchers, it is necessary to devise a mechanism for comparing different delineations both across method and across years. To this end, our analysis includes two techniques for representing change between delineations: one based on a binary 1/0 reporting whether a county is in the exact same grouping in two delineations, and another reporting the share of counties that are in the same cluster in both classifications. Defined mathematically, the two methods for scoring the 'fit' F for each county i between delineations x and y are calculated such that:

$$F_{ixy} = \text{if}\left(C_{ijx} = C_{ijy}\right) \text{then } 1, \text{ else } 0 \tag{2}$$

$$F_{ixy} = \frac{1}{2}\left(\frac{\sum_{j=1}^{J \in C_{ijx}} C_{ijx} = C_{ijy}}{J} + \frac{\sum_{j=1}^{J \in C_{ijy}} C_{ijy} = C_{ijx}}{J}\right), \tag{3}$$

where C indicates a vector of counties j assigned to the same cluster as i and the subscripts x and y denote the two delineations being compared. The test of equality in both equations is meant to return a 1 in the case of county j being in the same cluster as i in both delineations and 0 otherwise. Note that in Eq. 3 since i may be grouped with a different set of counties in x and y, we use the average of C_{ijx} in C_{ijy} and C_{ijy} in C_{ijx}. Both measures range between 0 (no congruence between clusters) and 1 (complete congruence between clusters. From these comparisons, we can generate an average score for the U.S. as well as choropleth maps showing the range of scores at the county level.

Calibration

With metrics in place for judging the quality of fit between two cluster delineations, it is straightforward to conduct an optimization procedure that tests different cutoff points for the clustering algorithm described previously. We can define an objective function that sums the scores for one of the two comparison methods offered in Eqs. 2 and 3 across all counties. A cutoff value that increases the correspondence of the two clusters under comparison will maximize this objective function. The optimization function using Eq. 3 was instrumental in identifying the .977 cutoff point we used for our replication without regional sub-groups.

Limitations of the Methods Described Here

The methods just described are useful in the sense that they offer a set of mechanisms for approximating the CZ designations developed in 1990 and 2000 using the approach documented by Tolbert and Killian. They offer a path forward to delineating clusters using the 2006–2010 ACS data, thereby extending the capacity of researchers to undertake analyses across a longer time period using more recent data. Perhaps most important, they provide a framework for comparing delineations using different methods or different years of data. One result of this comparison, however, is to shed some light on the limitations of the original methodology. While mitigating these limitations is beyond the scope of this paper, it is important that researchers using CZ data be aware of them.

The most significant limitation of the proportional flow methodology is that it privileges connections between large counties and small counties. Because of the denominator in Eq. 1, links between counties where one of the counties is much smaller than the other are understood as the most certain pairings for the purposes of cluster analysis. The denominator is based on the county with the smallest workforce of the pair, while the numerator is the sum of all connections between the two, so pairings with one small and one large county have the potential for a very small denominator and a substantial numerator making them stand out as points of strong connection in the context of the cluster analysis. One kind of connection that is under-valued in this representation would be two populous counties adjacent to one another and connected to many small counties on their periphery. The cluster analysis method employed here would put most of the weight on the connections with the various small counties and little weight on the connection between the two large counties. In practice, this configuration might reflect a metropolitan core that extends across two counties and our delineation would artificially split the core in two even as it recognized the core to periphery connections.

A second limitation of the methodology described here is that it works with average linkages, so that a candidate county's suitability for joining a cluster already made up of three counties is a reflection of the average connection it has with all three counties. If, however, the candidate county has a very strong connection with one county, but little or no connection with the other counties, it may not get selected for inclusion. Using the example of a large metropolitan area with a significant core and smaller periphery, this might result in portions of the periphery being peeled off inaccurately because they have little connection to other peripheral counties whose small size, as mentioned in the previous paragraph, would likely make them the first counties to join in a particular cluster.

A third limitation of the methodology employed here is that it does nothing to convey the differences across clusters in terms of goodness of fit, compactness, or any number of other possible measures of cluster quality. All we have is a sense that expert opinion viewed these cluster delineations to be reasonable in 1987 and our comparisons are based on congruence with that original determination. A more robust set of delineations would further unpack what we actually mean by "commuting zone" and consider additional metrics that capture the degree to which our clusters conform to this intended meaning rather than just how they conform to the original delineation. In this way, we could better ascertain whether the number, size, and configuration of our clusters are consistent with what we think we are describing in our analysis. Following the analysis portion of this paper, we offer some recommendations for how the limitations just presented might be mitigated.

Analysis

In this section, we compare the results of the several delineations described above and map the results of our comparisons. We also use our comparison framework for considering the changes in delineation over time both for the ERS delineations and for our replications. We conclude this section by offering a recommendation that researchers interested in continuity across decades utilize the approach relying on national-level data clustered with a cutoff point of .977 as it offers a very close approximation of the original delineation with a straightforward and parsimonious approach.

Side-by-side Comparison of Replication Strategies

Table 1 offers an assessment of the goodness of fit for our two methods of replicating the 1990 ERS delineations. Surprisingly, the attempt to exactly replicate Tolbert and Killian's method where we broke the country into overlapping sub-regions places the average county in a cluster with 90 % of the counties that were with it in the cluster assigned by Tolbert and Killian. This suggests that the method of adjudicating clusters placed into more than one cluster is not a close approximation of the expert knowledge approach undertaken in the original study. In contrast, the simpler "Full Country" replication where we did not break the country into sub-regions aligns with the original delineation about 94 % of the time.

We can gain a better understanding of where our replications succeed or do poorly by mapping the fit scores (generated as in Eq. 3) for each county comparing our Replication and the Full Country alternative with the 1990 ERS original delineation. In Fig. 1, we observe a significant portion of the Central or Midwest region to have relatively weak fit; these are the counties that were most likely to be assigned to multiple clusters as the overlap of sub-regions was most pronounced here. The relatively weak fit in this region suggests that our adjudication process was not terribly robust, and that any gains we might have made by running exactly the same cluster analysis as the original delineation were outweighed by the differences produced as clusters in the overlapping areas got sorted out. By way of comparison, Fig. 2 compares our full country delineation with the original ERS delineation with much better results. We can see in Fig. 2 that large swathes of the country are assigned to exactly the same cluster in both delineations. It is only in the Mid-Atlantic region and parts of the South where we have significant numbers of counties that have been assigned differently.

Table 1 Average similarity of county cluster assignments across delineations for 1990

	ERS 1990	Replication (%)	Full country (%)
ERS 1990	100.0		
Replication	89.7	100.0	
Full country	94.1	93.1	100.0

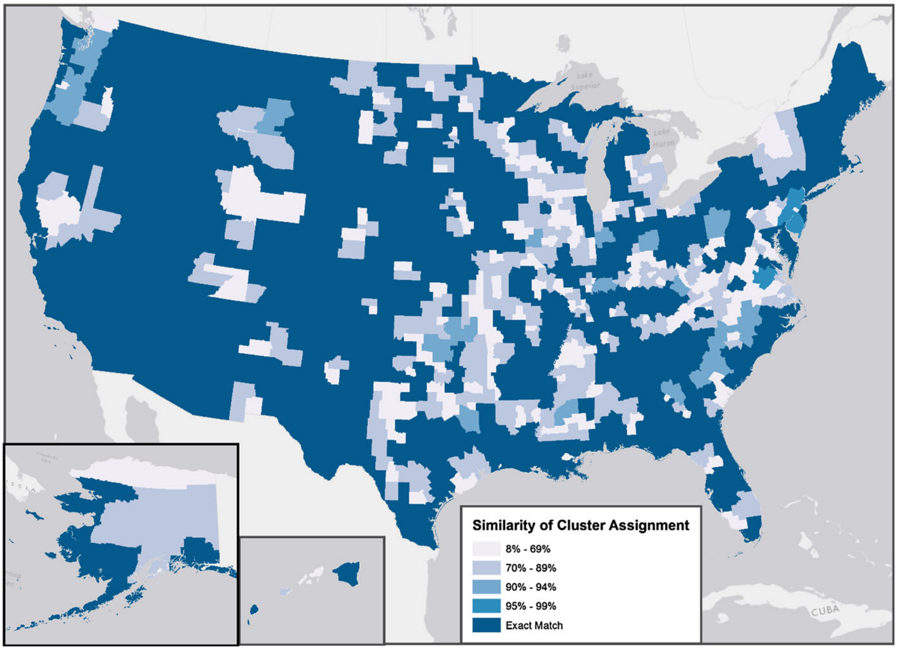

Fig. 1 Average fit in clusters 1990 replication with sub-regions versus ERS delineation

The side-by-side comparison indicates a strong reason to prefer the full country clustering method for both consistency and simplicity. Having made the case for this method, we now turn to understanding how CZ delineations have changed over time. To facilitate comparison, we exclude Alaska from our comparisons and aggregate several counties that changed during the time period. While there are reasons to prefer the full country clustering method, for the sake of completeness we provide CZ delineations for both the full country and the sub-regional methods in all three study years. These files will be appropriate for cross-sectional analysis as well as work that requires consistency with the earlier ERS delineations.

Comparing Change Over Time

A second key goal of this paper is to better understand how and where CZs have changed over time. In part, the location and extent of this change can help us better understand what CZs really are in terms of units of observation—changes should reflect in some way the larger changes we understand to be taking place in the spatial distribution of people and employment. So for example, we might expect CZs to reflect expanding metropolitan areas in the Southwest and shrinking metros in the Upper Midwest. Another reason to be interested in the degree of change in CZs over time is because this change should tell us how meaningful analyses that consider observations at multiple points in time are likely to be. We expect some degree of change in our CZs, but if that change is too great then it will be difficult to

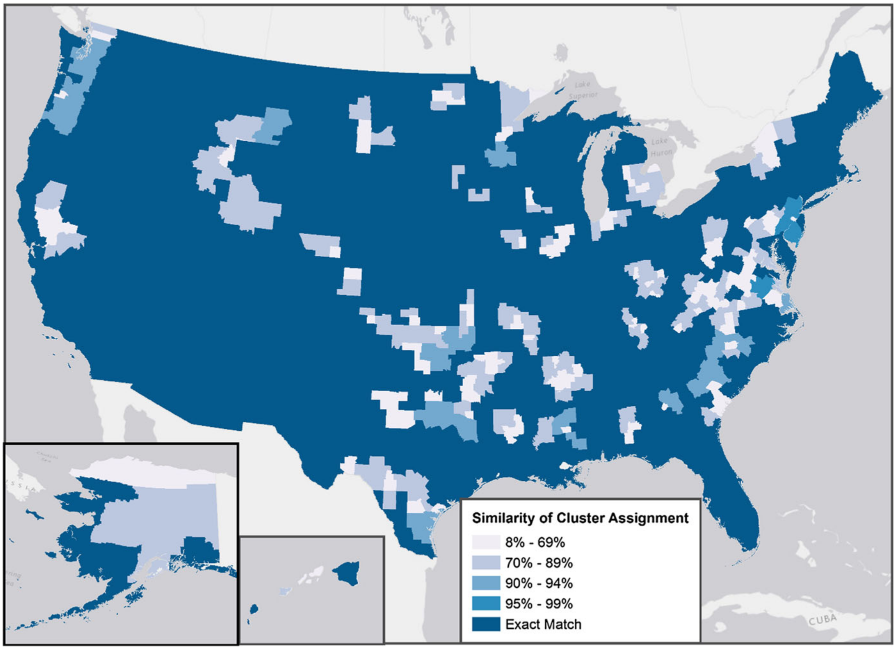

Fig. 2 Average fit in clusters 1990 full country versus ERS delineation

make meaningful comparisons over time. Finally, the stability of CZs over time is also relevant from the perspective of validating the methods for identifying CZs. While we do expect to see some change over time, too much change may be indicative of an arbitrary clustering mechanism that is not meaningfully capturing relationships among counties, but reflecting less important variation in the data.

Figure 3 documents the level of change between the 1990 and 2000 ERS CZ delineations. Clearly, there is a tremendous amount of change present in this figure. The average similarity score of 82.1 % means that the average county had only 82 % of the same counties in its cluster in 2000 that it had in 1990. These changes are widely dispersed with no immediately discernible spatial pattern. The extent of the changes, however, suggests that CZ definitions are not particularly stable over time as currently defined. In part this is a function of how the delineations were generated; the authors' discussions with ERS staff indicate that the 2000 delineations were not subject to the same expert review as in previous years.[10] More generally, it suggests that a more detailed understanding of how and why commuting zones change over time is overdue.

The changes present in comparisons of clusters across years using the full country clustering method yield strikingly similar results to those obtained in the ERS delineations. The average similarity for counties between 1990 and 2000 was virtually identical to the change registered in the ERS delineations (similar out to

[10] Authors' personal communication with ERS staff July 23rd, 2015.

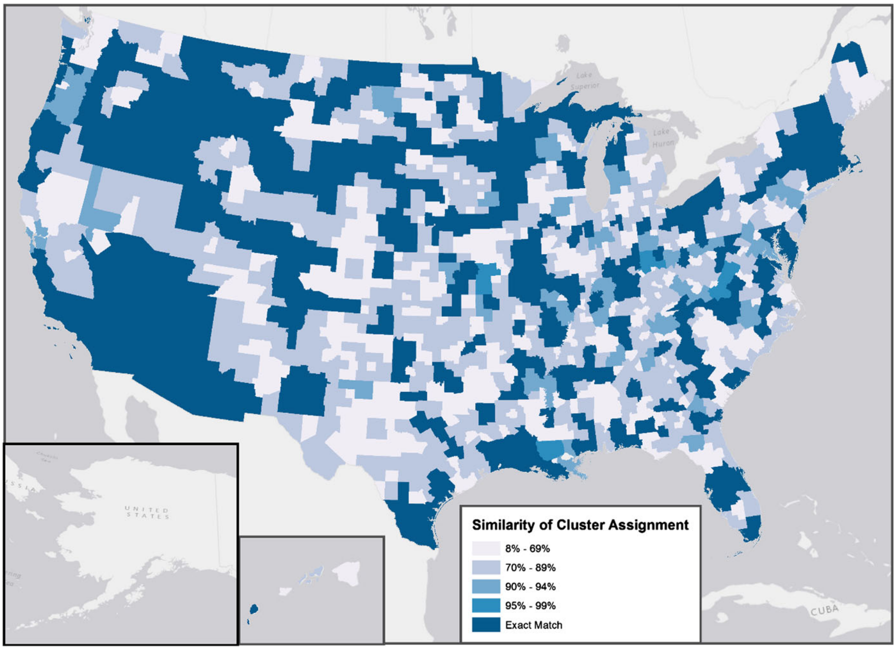

Fig. 3 Change in ERS cluster delineations 1990–2000

three decimal places in fact), dropping to 78.6 % when comparing clusters across the 1990–2010 period. While this represents a relatively small drop of only 3.5 % from the change between 1990 and 2000, a comparison of the 2000 and 2010 cluster delineations indicates a similarity of only 81.8 %. Clearly there is a great deal of shifting of cluster memberships between decades. Figure 4 shows the changes from 1990 to 2010 in our full country delineations and it looks similar enough to the ERS delineations to give us confidence that the full country method provides results that are consistent with those of ERS. More broadly, however, the overall level of change in cluster membership is cause for some concern with respect to the validity of these clusters if they shift so much between decades. Although we urge caution in conducting comparisons across decades due to this shift, we do provide an additional file with the counties and CZs aggregated and renamed to facilitate longitudinal comparisons.

Future Work

Careful scrutiny of the ERS clusters and of the methods we have designed to replicate the intent of those clusters suggests that further work is needed to revisit the original methodology and to update it so that it more certainly reflects both the stability and change of commute sheds in the United States. This is not just a matter of extending the original method as we have here, but of fundamentally rethinking how we designate CZs and how we capture both the stability and changing

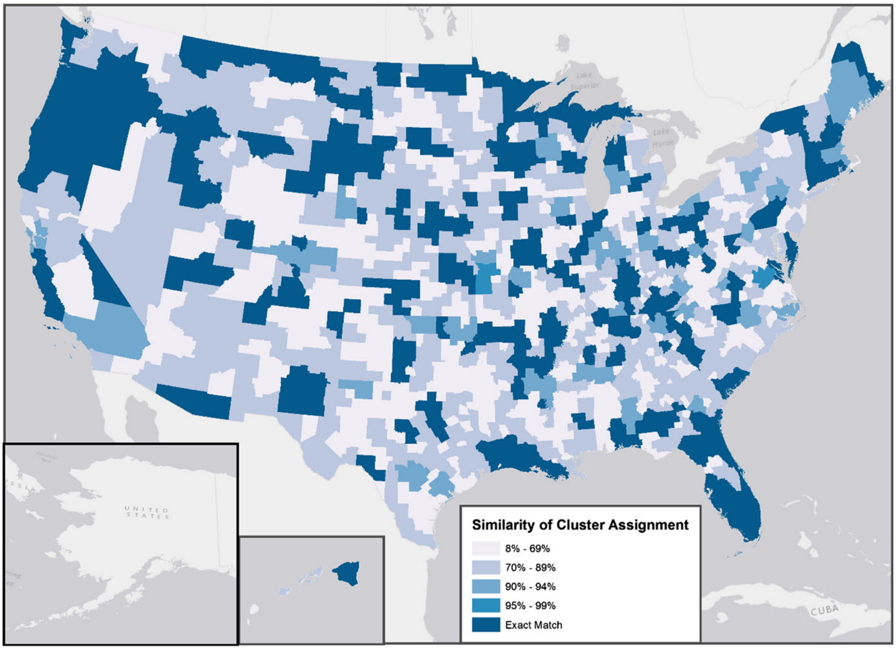

Fig. 4 Change in full country delineations 1990–2010

functionality of these zones. In this section, we briefly document some approaches that we feel will be fruitful in this regard that are part of an ongoing research project.

First, the data employed in our analyses could be improved, especially for rural areas. A primary reason for developing CZs is to document the connection between rural areas and urban cores; however, the thin coverage of rural areas in the ACS and the need to protect the privacy of ACS respondents in these areas means that there is tremendous uncertainty in the commuting flow estimates for less populous counties. Future work can benefit from more detailed geographic information available in the restricted files of the ACS and from the restricted use data available in the LEHD both available through Census Research Data Centers. Future work will need to, at a minimum, compare CZ delineations with and without these restricted data to ascertain the significance of the uncertainty about flows in rural areas.

Next, the proportional flow method of identifying links between counties has clear limitations as it favors uneven links between small and large counties at the expense of connections between pairs of counties of similar size. This drawback means that the methods employed in this paper may do a poor job of grouping core metropolitan counties together even as they are successful in capturing core periphery relationships. This limitation would appear to be increasingly important as the U.S. experiences continued growth in employment in suburban counties and in suburb-to-suburb commuting patterns.

To remedy the problem of similar-sized counties not being grouped in clusters, we propose a modification of a clustering method developed by Feser and colleagues for the purpose of identifying industry clusters (Feser 2005; Feser et al. 2008). The method proposed in Feser's (2005) paper generates clusters after considering the relative strength of four different kinds of relationships so that weakness in one kind of connection between counties may be offset by strength in another kind of connection. Applied to the case we are examining here, we could imagine the proportional flows of Eq. 1 constituting one measure of connectivity, while another measure could look at adjacency, similarity of size, and similarity of shares of flows internal, sent, and received. By combining multiple measures of connection as Feser does, we could build clusters that acknowledge different kinds of connections.

A second problem we see with the clustering methods undertaken to date is that they do little to convey the strength of connection between a county and its assigned cluster or the strength of a cluster with respect to its capture of nearby commuting flows. Clustering techniques simplify incredibly complex data relationships, but by their nature they make in/out assignments where connections are in fact much more nuanced. This is a major weakness of the method as we saw in our analysis of change over time. A cluster delineation changes in ways that appear arbitrary over time because some of the members assigned to a cluster may have been borderline assignments and slight changes in configurations of the data can lead to major reassignments of the clusters. To remedy this problem, we need a better way of representing both the strength of attachment between a county and its cluster and a measure of the quality of the cluster itself.

In practice, we already have partial solutions to the problems just described and the issue is that we need to incorporate these solutions into the clustering assignment process not just use them as post hoc measures of success. The "average distance" measure on which the clustering method is currently based (and which we exploited with limited success in our adjudication mechanism for our replication of the Tolbert and Killian method) is a good measure of strength of attachment but it needs to be applied in a way that recognizes a given county's attachment to the core of the cluster not just to the individual counties that make up the cluster. In simplest terms, average distance could be weighted by population size so that average distance to the largest counties in a cluster would be more important than the average distance to the smallest counties in a cluster. In practice, this could be as simple as using a different approach to clustering such as partitioning around medoids (Kaufman and Rousseeuw 2009) wherein clusters are represented by their center of mass rather than their component pieces. A second partial solution might involve greater incorporation of average distance into visualization and validation of clusters so that researchers could differentiate between changes over time that involve the shifting of weakly connected peripheral counties versus those that represent major reconfigurations of entire clusters. While the former is likely an accurate portrayal of the intent of CZs, the latter represents a likely flaw that would require rethinking of the method. That we do not employ these measures at present remains an important limitation of the efficacy of CZs.

While it is relatively straightforward to alter our clustering and visualization methods to differentiate between strong and weak connections between a county and its assigned cluster, incorporating measures of cluster strength into our analysis poses more significant problems. Conceptually, we can think of CZs as basins of attraction in which we expect most if not all commuter trips to take place within the zone. At present, we do not employ any measures that report on the degree to which a given county and CZ reflect what commute shed intuitively represents, the degree to which commuter flows are captured within the commuting zone. The absence of such a measure indicates that we are not completely sure that our theoretical conception of a CZ is matched in our data assignment. Certainly, we would expect CZs to be different in the densely populated East Coast with its multiple established metropolitan areas in close proximity than we would in Colorado or Montana where urban areas are more widely spaced. But how different are these CZs and should they even be characterized in the same way remains an open question because we do not have a clear idea of what constitutes a strong cluster or a weak one. We can imagine a clustering method that tries to optimize county assignment to clusters in a way that maximizes cluster commute capture and we are intrigued by the idea of a method that would delineate true commute sheds (say the maximum number of clusters for which at least 95 % of commutes are internal to the cluster); at present, we know of no existing methodology that would achieve this, but the development of such a technique is the subject of ongoing research.

A final issue with CZs as currently specified is that there is little work trying to understand the internal function of these zones. From everything we know about changing urban morphology, we would expect our CZs to have undergone significant internal change during the 20-year time period covered in the analysis presented here. We have more commuters living in the city and working in the suburbs and many more suburb-to-suburb commuters. There is evidence that the suburbs themselves are changing in terms of the socio-economic profile of residents, and the period from 1990 to 2010 saw significant growth in some cities and contraction in others. All of this change within CZs has gone largely unexplored and we feel that an adequate validation of CZs as units of observation must also include a better understanding of how these zones are changing internally.

In response to this last issue, we feel that future work needs to look at commuting at finer geographic scales and in the context of changing socio-economic characteristics of commuters. We are particularly optimistic about the opportunities afforded by the American Community Survey. ACS data provide the opportunity to capture seasonal employment trends and economic cycles while also shedding light on processes taking place within CZs as well as between them. While some of this information is publicly available (see reference to LODES and LEHD in Abowd et al. 2009), a thorough analysis will likely need to include restricted data available from the Census through one of its Research Data Centers. In the restricted data, the Census has commute flows geocoded down to the place or even block level and so changing internal commute patterns linked to a wide range of demographic variables is possible.

Conclusion

This paper offers some background on the development of revised CZs for 2010 and a set of methodologically consistent CZs for 1990 and 2000. It also proposes a set of tools for comparing delineations of CZs generated with different methods and in the context of change over time. The delineations and the tools to study them are made available for researchers interested in capturing a continuous representation of regional economies across the United States. In preparing this set of revised delineations, we have shed light on the history of these measures. In making our own delineations, we have provided a clearer view of how important expert opinion was to the assignment of counties to clusters in the original process and offered a method for assigning counties to CZs that is consistent with the intent of the original process without relying on expert opinion. Finally, we have documented a series of concerns we have with the methodology that underlies the delineations presented here. We feel that these delineations have value for the continuity they offer researchers, but future work can and should improve on the underlying decision rules so that outcomes conform more closely to our theoretical conception of what a CZ actually represents.

Acknowledgments This project has been supported by a Cooperative Agreement (No. 58-6000-4-0053) with the Economic Research Service, USDA. We also acknowledge assistance provided by the Population Research Institute at Penn State University, which was supported by an infrastructure grant by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R24-HD041025). Leif Jensen was supported by a USDA-funded Hatch Multistate Project W-3001, "The Great Recession, Its Aftermath, and Patterns of Rural and Small Town Demographic Change," administered through Penn State College of Agricultural Sciences Experiment Station Project Number PEN04504.

References

Abowd, J. J., Stephens, B. B., Vilhuber, L., Andersson, F., McKinney, K. L., Roemer, M., & Woodcock, S. (2009). The LEHD infrastructure files and the creation of the quarterly workforce indicators. In T. Dunne, J. B. Jensen, & M. Roberts (Eds.), *Producer Dynamics: New Evidence from Micro Data*. Chicago: University of Chicago Press.

Allegretto, S., Dube, A., & Reich, M. (2009). Spatial Heterogeneity and Minimum Wages: Employment etimates for teens using cross-state commuting zones. Institute for Research on Labor and Employment: Working Paper Series. http://escholarship.org/uc/item/1x99m65f

Armington, C., & Acs, Z. J. (2010). The determinants of regional variation in new firm formation. *Regional Studies, 36*(1), 33–45. doi:10.1080/00343400120099843.

Autor, D., & Dorn, D. (2009). This job is "Getting Old:" measuring changes in job opportunities using occupational age structure. *American Economic Review, 99*(2), 45–51.

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review, 103*(5), 1553–1597.

Cotter, D. A. (2002). Poor people in poor places: Local opportunity structures and household poverty. *Rural Sociology, 67*(4), 534–555.

Davis, E. E., Connolly, L. S., & Weber, B. A. (2003). Local Labor Market Conditions and the Jobless Poor: How much does local job growth help in rural areas. *Journal of Agricultural and Resource Economics*, *28*(3), 503–518.

Dorn, D. (2009). *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. St. Gallen: University of St. Gallen.

Dupor, B., McCrory, P. (2014) A cup runneth over: Fiscal policy spillovers from the 2009 recover act. Federal Reserve Bank of St. Louis: Working Paper Series. http://research.stlouisfed.org/wp/2014/2014-029.pdf

Economic Research Service. (2015). Commuting zones and labor market areas documentation. United States Department of Agriculture. Retrieved from http://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/documentation.aspx. Accessed date on 26 May 2015

Feser, E. J. (2003). What regions do rather than make: A proposed set of knowledge-based occupation clusters. *Urban Studies, 40*(10), 1937–1958. doi:10.1080/0042098032000116059.

Feser, E. J. (2005). *Benchmark value chain industry clusters for applied regional research*. Urbana-Champaign: Department of Urban and Regional Planning and Regional Economics Applications Laboratory (REAL), University of Illinois at Urbana-Champaign.

Feser, E. J., & Isserman, A. (2005). *Clusters and rural economies in economic and geographic space*. Urbana-Champaign: University of Illinois.

Feser, E. J., Renski, H., & Goldstein, H. (2008). Clusters and economic development outcomes: An analysis of the link between clustering and industry growth. *Economic Development Quarterly, 22*(4), 324–344. doi:10.1177/0891242408325419.

Fowler, C. S., & Kleit, R. G. (2013). The effects of industry clusters on the poverty rate. *Economic Geography, 90*(2), 129–154.

Fullerton, A. S., & Villemez, W. J. (2011). Why does the spatial agglomeration of firms benefit workers? Examining the role of organizational diversity in U.S. Industries and Labor Markets. *Social Forces, 89*(4), 1145–1164. doi:10.1093/sf/89.4.1145.

Gentleman, R., & Lang, D. T. (2007). Statistical analyses and reproducible research. *Journal of Computational and Graphical Statistics, 16*(1), 1–23.

Gibbs, R. M., & Bernat, G. A. (1980). Rural industry clusters raise local earnings. *Rural Development Perspectives, 12*(3), 18–25.

Graham, M. R., Kutzbach, M. J., & McKenzie, B. (2014). *Design comparison of LODES and ACS commuting data products* (No. CES 14-38). CES working papers. Washington D.C. doi:10.1017/CBO9781107415324.004

Henderson, J. R., & McNamara, K. T. (2000). The location of food manufacturing plant investments in corn belt counties. *Journal of Agricultural and Resource Economics, 25*(2), 680–697. doi:10.2307/40987084.

Hildner, K. F., Nichols, A., & Martin, S. (2015). *Methodology and assumptions for the mapping America' s futures project*. Washington D.C.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: An introduction to cluster analysis*. New York: John Wiley & Sons.

Killian, M. S., & Tolbert, C. M. (1993). Mapping social and economic space: The delineation of local labor market areas in the United States. In J. Singelmann & F. A. Deseran (Eds.), *Inequality in local labor markets* (pp. 69–79). Boulder, CO: Westview Press.

Lloyd, C., Shuttleworth, I., & Wong, D. (Eds.). (2014). *Social-spatial segregation: Concepts, processes and outcomes*. Bristol: Policy Press.

Martin, S., Astone, N. M., Peters, H. E., Pendall, R., Nichols, A., Hildner, K. F., & Stolte, A. (2015). *Evolving patterns in diversity*. Washington D.C.

McKenzie, B. S. (2013). *County-to-County Commuting Flows: 2006-2010*. US: Census Bureau.

Minnesota Population Center. (2011). *National historical geographic information system version 2.0*. Minneapolis, MN: University of Minnesota.

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives, 25*(3), 173–196. doi:10.1257/jep.25.3.173.

Nichols, A., Martin, S., Astone, N., & Peters, H. (2015). *The labor force in an aging and growing America*. Washington, DC: The Urban Institute.

Pendall, R., Martin, S., Astone, N., & Nichols, A. (2015). *Scenarios for regional growth from 2010 to 2030*. Washington, DC: The Urban Institute.

Spear, B. D. (2011). *Improving employment data for transportation planning NCHRP 08-36, Task 098*. Cambridge, Massachusetts.

Tickamyer, A., & Bokemeier, J. (1987). Sex differences in labor-market experiences. *Rural Sociology, 53*(2), 166–189.

Tolbert, C., Blanchard, T., & Irwin, M. (2009). Measuring migration: Profiling residential mobility across two decades. *Journal of Applied Social Science, 3*(2), 24–38.

Tolbert, C., & Killian, M. (1987). *Labor market areas for the United States*. Washington, DC: US Department of Agriculture.

Tolbert, C., & Sizer, M. (1996). *U.S. commuting zones and labor market areas: A 1990 update*. Washington, DC: US Department of Agriculture.

United States Census Bureau. (2010a). Decennial Census Questionnaires for 1990 and 2000.

United States Census Bureau (2010b). Decennial Census of the Unites States.

Voss, P. R. (2007). Demography as a spatial social science. *Population Research and Policy Review, 26*(5–6), 457–476. doi:10.1007/s11113-007-9047-4.

Zolnik, E. J. (2010). The geographic distribution of U.S. unemployment by gender. *Economic Development Quarterly, 25*(1), 91–103. doi:10.1177/0891242410386592.