

Supplemental information:

Justification for Descriptive Measures

Completeness plays an important role in labor-shed delineation and represents a major point of difference across extant delineations. The labor market concept is easy to explain when considering the array of counties surrounding a large metropolitan area but can become strained in remote rural areas. For this reason some delineations, like OMB's CBSAs, only include counties that are demonstrably attached to a specific metropolitan area with strong commuting flow relationships. In contrast, the ERS Commuting Zones reflect a greater interest in connectivity for rural areas and include counties that may have a very slight connection. The intent of an analysis should determine whether completeness and inclusion of rural populations is of greater importance than conformance to the labor market concept. Decision metrics included here describe the number of counties and share of the population covered.

Size (referring to physical area, number of sub-units, and population) *and shape* are relevant metrics for judging the quality of a labor market delineation and also convey important information about the scale of observations an analyst will be using. Metrics describing the size and skew (kurtosis) typical to a delineation offer a comparison of both the ‘typical’ labor market and the extremes of the delineation. While we would expect labor markets to vary significantly in size both for historical reasons and due to the functioning of economic systems (e.g. Zipf, 1949), maintaining some degree of homogeneity among observations is desirable in many analytic contexts.

Compactness and *Contiguity*, here presented as aspects of shape, have a basis in previous work on regionalization problems (Duque et al., 2007). Compactness typically refers to the ratio

of the area to the perimeter. The Iso-Perimeter Quotient (IPQ) is a standard measure for compactness (for a review see Li et al., 2013). The IPQ is defined as:

$$IPQ = \frac{4\pi Area}{Perimeter^2}$$

The IPQ ranges from 0 to 1 with one indicating that the shape is perfectly compact (i.e. circular). Beyond flagging delineations that result in particularly unusual combinations of counties, compactness is not a particularly useful measure in this context as it is largely determined by the shape of the constituent counties and there is little variation among the finalized delineations presented for comparison. Contiguity plays a similar role to that of compactness, as conceptually, labor-sheds are typically contiguous, but the finished delineations compared here are all composed of contiguous labor-sheds so there is no variation on this variable.

Spatial variation of fit measures represents a crucial metric for describing delineations as the clustering of similar fit scores may signal spatial variation in the quality of observations that might be expected to bias results of analyses conducted with these units. We compute Moran's *I* (Moran, 1950) as an indicator of clustering and locate it for convenience among descriptive statistics, however, this metric is calculated for a range of fit statistics described below and describes the degree of spatial variation in fit across each of our key definitional areas. In each case *I* is calculated using the delineation units (e.g. Economic Areas, Commuting Zones) and 'neighbor' is based on the nearest five observation centroids. The nearest neighbor calculation is preferred to contiguity as a measure of neighborhood in this case because it works for both complete (Bureau of Economic Analysis, BEA and ERS) and partial (CBSA, Tong and Plane) delineations.

Justification of measures of Core focus

Metropolitan areas are understood to be the engines of most regional economies (Giuliani, 2007; Katz and Bradley, 2013; Krugman, 1991; Partridge et al., 2007; Sole and Viladecans-marsal, 2004). While the definition of metropolitan area has changed considerably over time (appropriately given the changed spatial structure and morphology of urban areas), having *metropolitan areas split into multiple labor markets* does not make sense for most applications. In some places this may be a necessary outcome of using counties as sub-regions; as in the Northeast where metropolitan area “commuting fields” are substantially overlapping and do not conform to county boundaries (Plane, 1981). In general, however, it will be desirable to minimize the number of metro areas that are split. Here we use decade-appropriate metropolitan boundaries and identify the number of metropolitan areas that contain counties assigned to more than one labor market.

In addition to assessing the degree to which metropolitan areas are split, we also identify instances of labor markets that do not have *at least one core county* associated with them. Core counties are defined by OMB as having an urbanized area (itself defined at the tract level) of at least 10,000 people. For some applications, the idea of a labor market reasonably requires a concentration of population and economic activity, and the absence of counties with these designations will represent a significant limitation.

Finally, we include a measure of the *average share of county residents in a labor market who work in a core county of that labor market*. This measure indicates the degree to which core counties are the employment engine for the labor market and will be high for labor markets with large core counties, or counties that draw in large numbers of residents from peripheral regions. Note that by averaging across counties in the labor market we dramatically increase the role of

low-population counties that might otherwise be masked by larger counties in the labor market. For completeness we also calculate the share of each county's workforce that resides in a core county in the labor market (not reported) and report the sum of the two measures (share of county residents working in a core county in the labor market plus share of county workforce residing in a core county in the labor market). This measure is reported because it mimics the metric used by OMB (2010) to determine whether a county should be attached to a CBSA. Current OMB standards require this measure to be at least 25% (although earlier versions of the standard required only 15% with local support for inclusion in the CBSA).ⁱ

Justification of measures of Connection

One reason that researchers employ labor markets is to reflect units where labor conditions might be expected to move together. Within a labor market we would expect unemployment, wages, and similar measures to converge across sub-regions as workers within the labor market could be expected to compete for the same jobs. While more measures of connectivity are certainly possible, for simplicity we focus here on correlations among counties in the average wage. Specifically, we adopt the method employed by Foote et al. (2016, page 17, 2017) and calculate the population-weighted average pairwise correlation in seasonally-adjusted average wages over a six-year time period among all pairs of counties in the same labor market.ⁱⁱ

For a given labor market C the average pairwise correlation in wages p_C is defined such that:

$$p_C = \frac{1}{2N} \sum_{i \in C} \sum_{j \in C} w_{ij} p_{ij}$$

Where the weights in w_{ij} are defined by lf_i , the size of the labor force in each county such that:

$$w_{ij} = \frac{lf_i + lf_j}{2 \sum_{k \in C} lf_k}$$

and p_{ij} is based on six years of data beginning with the decade for which the measure is defined (e.g. 1990 through 1995 for 1990 delineations). In cases where a labor market has just a single county the correlation is set to 1.ⁱⁱⁱ

Justification of measures of Containment

The final set of criteria we examine measure containment of the employed population, with its basis in the natural phenomenon of a watershed where every drop of rain that falls within a watershed remains there until evaporation. The idea of containment is a foundational component of delineating labor markets (Coombes et al., 1986, page 944) and is crucial for our definition of labor-shed. Having a high degree of containment means that there should be minimal interactions between labor markets (Cörvers and Hensen, 2003; Smart, 1974). Containment can be thought of in two ways: home-containment and work-containment. Home-containment is the share of the resident population who also work in the labor market. Work-containment is the share of jobs in the labor market filled by residents of the labor market.

The final measure included here, the share of the U.S. population whose commute is contained within their assigned labor market, is a single overview value for the entire delineation. Whereas the previous measure averages containment across all observations in a given delineation (giving a good sense of the average quality of observations, but making no distinction between heavily populated and lightly populated labor markets), this metric looks at the population of the country as a whole. It is included here as it was a key metric used by Tong and Plane (2014) to justify their delineation as an improvement over CBSA's.

References

Coombes M G, Green A E, Openshaw S, 1986, "An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-to-Work Areas Revision in Britain" *Journal of the Operational Research Society* **37**(10) 943–953, <https://www.tandfonline.com/doi/full/10.1057/jors.1986.163>

- Cörvers F, Hensen M, 2003, "The regionalization of labour markets by modelling commuting behaviour", in *43rd Congress of the European Regional Science Association: "Peripheries, Centres, and Spatial Development in the New Europe"* (Louvain-la-Neuve: European Regional Science Association (ERSA), Jyväskylä, Finland),
<https://www.econstor.eu/handle/10419/116031>
- Duque J C, Ramos R, Suriñach J, 2007, "Supervised Regionalization Methods: A Survey" *International Regional Science Review* **30**(3) 195–220,
<http://journals.sagepub.com/doi/10.1177/0160017607301605>
- Foote A, Kutzbach M, Vilhuber L, 2016, "Driving past commuting zones: Re-examining local labor market definitions" *2017 ASSA Annual Meeting* 1–26,
<https://www.aeaweb.org/conference/2017/preliminary/paper/thT52i7D>
- Foote A, Kutzbach M, Vilhuber L, 2017, "Recalculating - How Uncertainty in Local Labor Market Definitions Affects Empirical Findings"
- Giuliani E, 2007, "The selective nature of knowledge networks in clusters: evidence from the wine industry" *Journal of Economic Geography* **7**(2) 139–168,
<https://academic.oup.com/joeg/article-lookup/doi/10.1093/jeg/lbl014>
- Katz B, Bradley J, 2013 *The metropolitan revolution: How cities and metros are fixing our broken politics and fragile economy* (Brookings Institution Press, Washington D.C.)
- Krugman P, 1991, "Increasing Returns and Economic Geography" *The Journal of Political Economy* **99**(3) 483–489
- Li W, Goodchild M F, Church R, 2013, "An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems" *International Journal of Geographical Information Science* **27**(6) 1227–1250,
<http://www.tandfonline.com/doi/abs/10.1080/13658816.2012.752093>
- Moran P A P, 1950, "Notes on continuous stochastic phenomena" *Biometrika* **37**(1) 17–23,
https://www.jstor.org/stable/2332142?casa_token=1AsEl9k_HskAAAAA:GU4r6tftYQa_ImqD6Xpe-TU4x7fO-EC7gfjT3Fi_ux5jUIaQGcNw6bnsTZ7Bp6Ct048nOw6Z_hwsV2aeibUvKNVTC0G6m9quCsvsoytHfc9IV0Uq5g
- Office of Management and Budget, 2010, "2010 Standards for delineating metropolitan and micropolitan statistical areas" 37245–37252,
<https://www.federalregister.gov/documents/2010/06/28/2010-15605/2010-standards-for-delineating-metropolitan-and-micropolitan-statistical-areas>
- Partridge M, Olfert M R, Alasia A, 2007, "Canadian Cities As Regional Engines of Growth: Agglomeration and Amenities (Les villes canadiennes en tant qu'engins de croissance: agglomération ou commodités?)" *The Canadian Journal of Economics / Revue canadienne d'Economique* **40**(1) 39–68, <http://www.jstor.org.ezproxy.library.ubc.ca/stable/4620596>
- Plane D A, 1981, "The geography of urban commuting fields: Some empirical evidence from new England" *Professional Geographer* **33**(2) 182–188,
<http://onlinelibrary.wiley.com/doi/10.1111/j.0033-0124.1981.00182.x/full>
- Smart M ., 1974, "Labour market areas: Uses and definition" *Progress in Planning* **2** 239–353,
<https://www.sciencedirect.com/science/article/pii/0305900674900087>
- Sole A, Viladecans-marsal E, 2004, "CENTRAL CITIES AS ENGINES OF METROPOLITAN AREA GROWTH *" **44**(2) 321–350
- Zipf G K, 1949 *Human behavior and the principle of least effort* (Addison-Wesley Publishing Company, Oxford, UK)

ⁱ Our analysis found several counties that were included in 2010 OMB CBSA definitions that did not meet the thresholds for inclusion. Given the rolling nature of the ACS-based commuting calculations and some uncertainty as to exactly what year and version of the ACS was used variations significant enough to alter a county's inclusion or exclusion are inevitable. In fact, the Margins of Error placed on ACS data mean that significant differences year to year are both possible and likely. We found counties with as little as 2% of their residents working in a core area of their CBSA and none of their local workforce commuting in from a core county in the CBSA.

ⁱⁱ Foote et al. (2016) is a preliminary draft and cited with permission of the authors. Foote et al. (2017) is the finished paper but omits the crucial pairwise correlation formula referenced here. We cite both for completeness.

ⁱⁱⁱ The decision to assign 1 to single county labor markets favors delineations with a larger number of single-county labor markets. This is not significant for most delineations, but is highly significant for CBSA's especially micropolitan CBSA's. To account for this we include information on how many single county labor markets there are in each delineation and calculate the mean wage correlation with and without single county labor markets.