Where is the City’s Center?  
Five Measures of Central Location

Matthew J. Holian  
San Jose State University

Abstract

In this article, the author describes and evaluates five databases that provide different measures of city center location. He identifies research that has used each database, calculates average distances between the locations, presents case studies for two cities, and provides suggestions for analysts searching for appropriate measures of city center location. Among the findings are that the location of a city’s city hall is a better proxy for the location of a city’s central business district (CBD) than are other measures in current use.
Introduction

Researchers across many disciplines need city center location data to identify such places as central business districts (CBDs). The notion of a “city center”—as distinct from the “central city” of a metropolitan statistical area (MSA)—has different meanings depending on the context. One may envision the peak of a city skyline (often in financial centers); a historic transportation hub, such as Grand Central Station in New York City (Schelling, 1960); or a location defined on objective criteria, such as population or employment density (Taubenböck et al., 2013). Other meanings of the term city center could be possible. The concept of a CBD, however, is widely used in research.

In this article, I present five measures used in urban economics and related research to identify city center locations. I describe strengths and weaknesses of the measures and present an analysis comparing the five databases, four of which cover U.S. MSA centers and a fifth that includes all census geographies. In keeping with the applied focus of the SpAM department of this journal, I also describe my methods so that others can adopt the techniques. I conclude the article with suggestions for finding appropriate measures of the socioeconomic center of a city.

Five Measures of Central City Points

A point—that is, a location—is represented on a map by geographic coordinates, which give the latitude and longitude of the “geocoded” point. At least five different measures have been used by researchers to identify city center locations. In this section, I describe those measures and outline their differences.

The first measure is from Central Business Districts: 1982 Census of Retail Trade (hereafter, 1982 Census), which was the last time the U.S. Census Bureau attempted to identify CBDs, and it did so for 455 cities defined by existing census tract boundaries. Kneebone (2013: 3) notes that “Though dated, the 1982 CBDs represent the last systematic identification of business districts at the national scale. Furthermore, the 1982 CBDs continue to exhibit significant overlap with the densest job centers in the nation’s major metro areas.” Despite the strong conceptual appeal of the 1982 Census measure, one limitation for a researcher looking for geographic points is that those tract-based definitions are actually areas.

Researchers working at the Federal Reserve have converted those areas to points and made the data available on the Internet. Although I refer to this measure as the 1982 Census measure, the conversion of areas to points was performed by Fee and Hartley (2013), who required CBD points for 385 MSAs. They describe their methodology on pages 47 to 48 of their study:

---

1 A new and promising approach uses machine learning techniques along with both subjective assessment and objective criteria to determine CBD points (Brown et al., 2017).
2 This is my manual count from the PDF file available at the Internet archive: https://web.archive.org/web/20070221191519/http://www.census.gov:80/geo/tiger/cbdct.pdf. This file lists census tracts identified by the 1982 study as a “central business district.” See Brown et al. (2017: 6) for more details on the 1982 Census delineation of CBDs.
3 These data can be accessed at http://www.danielaaronhartley.com/msas_with_central_city_cbds.csv.
We identify the latitude and longitude of the CBD by taking the collection of census tracts listed in the 1982 Census of Retail Trade for the central city of the MSA (the city in the MSA with the largest population) and finding the centroid of that cluster of census tracts. We identify the CBD latitude and longitude for 268 MSAs in this manner.

Note that because the 1982 Census measure is available for only 455 cities, Fee and Hartley (2013) were able to merge these data to only an N=268 subset of their sample of contemporary MSAs. Their solution to this problem, described by the authors on pages 47 to 48, takes us to a second measure of central city location:

For the remaining 117 MSAs, whose central city was not listed in the 1982 Census of Retail Trade, we use the latitude and longitude found by geocoding the MSAs central city found using the ArcGIS 10.0 North American Geocoding Service. ArcGIS returns points that are, on average, very close to the CBDs from the Census of Retail Trade; for the 268 cities for which we have both, the mean distance between the two is 0.39 miles.

One area of interest for Fee and Hartley (2013) was measuring the effect of so-called agglomeration economies, which refers to greater invention, innovation, and enhanced productivity because workers in the same industry live, work, or otherwise interact in close proximity. Fee and Hartley (2013) found an average distance of 0.39 miles between the 1982 Census and the ArcGIS measures. The authors concluded that the measurement error was likely small for most industries because Elvery and Sveikauskas (2010) suggest in their work that distances of between 5 and 25 miles may be relevant for many industries. They discuss the possibility, however, that relevant distances for agglomeration effects may be shorter in industries with more highly skilled workers.

A third measure, similar to the ArcGIS measure, was developed as part of the study by Holian and Kahn (2015), who used Google Earth to geocode the principal city of each of 366 MSAs by recording the latitude and longitude returned—the location where the map view centered—from a city name search. One of the applications in which Holian and Kahn (2015) were interested was whether households living near the city center drive less on average than those who live farther away. Given that the average metropolitan U.S. household lived about 20 miles from their city’s downtown area, measurement error—if on the scale of that reported by Fee and Hartley (2013)—would not seem to present a major threat to the internal validity of the analysis. Whether the Google Earth measure is close to the 1982 Census measure is unknown, but this topic is addressed in the next section.

---

4 ArcGIS is a prominent Geographic Information System (GIS) software package by Environmental Systems Research Institute (ESRI). ESRI provides GIS data files for use in ArcGIS. Fee and Hartley (2013) used these data to create CBD locations, but the authors do not describe how they determined the point location.

5 This possibility was raised by Fee and Hartley (2013) in endnote 2: 262.

6 These data are available to download on my blog: http://mattholian.blogspot.com/2013/05/central-business-district-geocodes.html.

7 In their initial study, Holian and Kahn (2015) recognized the possibility of measurement error but noted that errors of up to a mile would likely not present a major challenge given the nature of the analysis, and that all points in a sample of points they visually inspected seemed to be reasonable approximations of CBD location.
Like the ArcGIS measure, the Google Earth measure is also something of a “black box” in the sense that the variables in the Google algorithm that determined the locations are unknown. Nevertheless, this measure has been used or discussed in articles, including Anenberg and Kung (2018) and Gardner and Hendrickson (2018); Ph.D. dissertations, including Molnar (2014), Resseger (2014), and Su (2018); Federal Reserve Bank publications (Rappaport, 2017, 2016, and 2014); and working papers (Couture et al., 2018; Couture and Handbury, 2017; Hamilton and Dourado, 2018).

A fourth measure was developed by Wilson and colleagues (2012), who provide geocodes based on the address of each city's city hall. These data are from a worksheet in the Chapter 1 spreadsheet file associated with their study. A principal advantage of these measures is that they are conceptually clear, similar to the centers from the 1982 Census. Although the concept of a political center is distinct from that of a business center, one can reasonably expect these political centers to be near the CBDs. This city hall measure provides geocodes for the principal city in each of 368 MSAs. Examples of studies that have used or discussed location information from this source include Dascher (2018); Hall, Palsson, and Price (2018); and Liu (2013).

Finally, a fifth measure of city center used in research is the U.S. Census Gazetteer. These center points are based on the interpolated latitude and longitude associated with the Census Bureau’s TIGER/Line Shapefiles. The Gazetteer files provide population, land area, and “representative latitude and longitude coordinates” for all census geographies, including tracts, counties, and cities. Documentation explaining how these points were selected or created is unavailable. This measure is similar to Google Earth and ArcGIS with regard to being a black box in creating the central point; however, the Gazetteer is by far the most comprehensive measure I discuss here, which is its principal virtue and can be readily accessed for all census geographies. Holian and Kahn (2015); James and Aadland (2011); and Nelson, Uwasu, and Polasky (2007) provide examples of research use location information from this source.

Analysis

In the preceding section, I described conceptual strengths and weaknesses but posted several questions about their accuracy in representing a city center. This section provides a quantitative assessment of the measures and offers some answers as to their accuracy.

To calculate the distance between two points, I use a formula called the spherical law of cosines. This formula takes into account the spherical shape of the Earth and is accurate over distances
Where is the City’s Center? Five Measures of Central Location

The results in exhibit 1 show the City Hall measure to be closest to the 1982 Census measure, at 0.52 kilometers (km), and the Gazetteer measure is farthest from it, at 3.68 km. The mean distance between the ArcGIS and 1982 Census measures used by Fee and Hartley (2013) is 0.64 kilometers, which is extremely close to the mean distance of 0.63 km (or 0.39 miles) that they reported. Of note, I find nearly the same result as Fee and Hartley, even though my sample is slightly different, with N=256 versus N=268. Sample sizes for each mean distance measure in exhibit 1 are shown in exhibit 2.

I now offer a more detailed analysis of the measures. Exhibit 2 shows summary statistics for each of the 10 distances. In addition to average distance, exhibit 2 reports the standard deviation of the distances. These data are useful to report because a measure of location that is farther from the CBD on average but that has a smaller standard deviation may be more desirable for some purposes than a measure that is closer on average but for which the possibility of large errors exists. In addition, exhibit 2 reports the number of cities (N) and the median, minimum, and maximum distances to provide fuller insight into the characteristics of the distances between each CBD measure.14

---

13 This file and accompanying tutorial analysis can be downloaded from http://mattholian.blogspot.com/2017/03/using-spreadsheet-to-calculate-distance.html. For more information on the formula, see http://www.movable-type.co.uk/scripts/latlong.html. I have also written a blog article (Holian, 2017) that contains instructions for using that file, sample data, and a tutorial analysis. My “distance calculator” spreadsheet also contains a worksheet with additional instructions. Although I provide a lot of instructions, the spreadsheet is actually fairly easy to use. There are a few helpful practices that users would learn themselves after using it a few times, but I have tried to spell out all of them in the blog article so that others will not have to reinvent the wheel.

14 I did not include an estimate of the standard error (SE) of the means in repeated sampling, but that value can be readily calculated based on the formula SE=SD/√N. The test statistic for a test for which the mean distance is zero is then the mean divided by the SE. All of the mean distances are statistically significant at zero at conventional levels.
### Exhibit 2

#### Summary Statistics for 10 Distance Measures

<table>
<thead>
<tr>
<th>Distance Measurement</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982Cen to ArcGIS</td>
<td>256</td>
<td>0.64</td>
<td>0.42</td>
<td>0.78</td>
<td>0.01</td>
<td>6.85</td>
</tr>
<tr>
<td>1982Cen to Google Earth</td>
<td>256</td>
<td>0.97</td>
<td>0.49</td>
<td>1.53</td>
<td>0.02</td>
<td>12.62</td>
</tr>
<tr>
<td>1982Cen to City Hall</td>
<td>256</td>
<td>0.52</td>
<td>0.36</td>
<td>0.77</td>
<td>0.02</td>
<td>6.93</td>
</tr>
<tr>
<td>1982Cen to Gazetteer</td>
<td>256</td>
<td>3.68</td>
<td>2.51</td>
<td>4.65</td>
<td>0.06</td>
<td>55.26</td>
</tr>
<tr>
<td>ArcGIS to Google Earth</td>
<td>361</td>
<td>0.98</td>
<td>0.50</td>
<td>1.60</td>
<td>0.00</td>
<td>13.07</td>
</tr>
<tr>
<td>ArcGIS to City Hall</td>
<td>361</td>
<td>0.48</td>
<td>0.30</td>
<td>0.65</td>
<td>0.01</td>
<td>5.72</td>
</tr>
<tr>
<td>ArcGIS to Gazetteer</td>
<td>361</td>
<td>3.36</td>
<td>2.37</td>
<td>4.39</td>
<td>0.12</td>
<td>54.14</td>
</tr>
<tr>
<td>City Hall to Google Earth</td>
<td>365</td>
<td>0.94</td>
<td>0.45</td>
<td>1.65</td>
<td>0.01</td>
<td>13.40</td>
</tr>
<tr>
<td>City Hall to Gazetteer</td>
<td>365</td>
<td>3.44</td>
<td>2.40</td>
<td>4.46</td>
<td>0.06</td>
<td>54.27</td>
</tr>
<tr>
<td>Google Earth to Gazetteer</td>
<td>366</td>
<td>3.50</td>
<td>2.35</td>
<td>4.95</td>
<td>0.07</td>
<td>54.14</td>
</tr>
</tbody>
</table>

1982Cen = 1982 Census of Retail Trade. SD = standard deviation.

The ArcGIS measure is 0.64 km from the 1982 Census measure on average, whereas the City Hall measure is only 0.52 km away on average. That difference is a relatively small 0.12 km, but it suggests, for example, that because the City Hall measure is closer to the 1982 Census measure on average, studies such as Fee and Hartley (2013) and Holian and Kahn (2015) would have been better off using the City Hall measure from Wilson et al. (2012) instead of the ArcGIS or Google Earth measures. Is the difference large enough to be statistically significant? The value of the difference between means test statistic is 1.71 (p=0.087), meaning that the difference is significant at the 10-percent level but not at the more conventional 5-percent level. The standard deviation and maximum distances are very close for both the 1982 Census-to-City Hall and the 1982 Census-to-ArcGIS distances. Finally, in addition to the Gazetteer measure having the highest average distances to all other central location measures, the maximum distance is much higher in the Gazetteer measure and in one case is more than 54 km away from the City Hall measure, which is in the San Francisco MSA and is discussed in the next section.

---

15 See McDonald (2014) for a discussion about this test; the author of that article has also created a spreadsheet tool that implements this test, which can be downloaded from that page. This test is conceptually different from what is probably the more familiar difference-in-means hypothesis test.
Application

This section presents qualitative analysis for two MSAs as case studies, starting with New Orleans. The Google Earth measure is not the spot that most people would identify as the center of the city, but it is close to typical locations that people might cite. For example, the Google Earth measure locates the city center at slightly less than 1 mile north of Jackson Square, a central park in the city's historic French Quarter. That difference may or may not be a problem for some analyses, such as housing price change away from the city center, but that distance could be a large enough error to be a problem for other applications, such as in measuring the diffusion of ideas. For example, if workers on their lunch break are not able to walk much more than a mile from their office, measurement error on this scale could severely limit opportunities for informal exchange of ideas.

The New Orleans City Hall location is within the area that is known locally as the CBD, but it is toward the outskirts of that district, near the Superdome and Interstate 10. The ArcGIS measure is one large city block from the City Hall measure. Although these three may be close for many applications, the 1982 Census measure is the best of the five because it is firmly centered in the neighborhood identified by local residents as the CBD and is only a regular city block from the French Quarter. Meanwhile, the Gazetteer measure is the worst of the five; it locates the center in a wildlife refuge more than 10 miles away from the CBD and the French Quarter. Exhibit 3 shows geocodes for all five measures, exhibit 4 presents all five points, and exhibit 5 shows the four nearby points presented in greater detail.

### Exhibit 3

**Five Measures of Central Points for Two Cities**

<table>
<thead>
<tr>
<th>Measure</th>
<th>New Orleans Latitude</th>
<th>New Orleans Longitude</th>
<th>San Francisco Latitude</th>
<th>San Francisco Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Google Earth</td>
<td>29.964722</td>
<td>− 90.070556</td>
<td>37.774930</td>
<td>− 122.419416</td>
</tr>
<tr>
<td>2 City Hall</td>
<td>29.952403</td>
<td>− 90.076487</td>
<td>37.778503</td>
<td>− 122.418307</td>
</tr>
<tr>
<td>3 Gazetteer</td>
<td>30.068636</td>
<td>− 89.939007</td>
<td>37.727239</td>
<td>− 123.032229</td>
</tr>
<tr>
<td>4 Census 1982</td>
<td>29.952499</td>
<td>− 90.070801</td>
<td>37.785702</td>
<td>− 122.408000</td>
</tr>
<tr>
<td>5 ArcGIS</td>
<td>29.953701</td>
<td>− 90.077751</td>
<td>37.777122</td>
<td>− 122.419639</td>
</tr>
</tbody>
</table>

As another illustrative example with these measures, consider San Francisco, where the Gazetteer point is on the Farallon Islands, which are in the Pacific Ocean and are visible from the mainland only on the clearest of days. The City Hall, Google Earth, and ArcGIS measures, however, are within one to four blocks of each other. Despite the near consensus reached by those three measures, other contender definitions for the center of San Francisco—the historic transportation center (the Ferry building), financial district (FiDi), or shopping district (Union Square)—are several kilometers away from city hall. The City Hall and Google Earth measures are only a few kilometers away from each of those points, whereas the Gazetteer measure is more than 50 km away. Finally, and as in the case of New Orleans, the 1982 Census measure seems most appropriate for the majority of applications requiring a socioeconomic measure of centrality: it is 2 blocks from the Union Square shopping district, which is between the Financial District, the Ferry Building, and city hall.
Exhibit 5
Four Measures of Central Location in New Orleans, Detail Area

Notes: 1. Google Earth. 2. City Hall. 4. Census 1982. 5. ArcGIS. Plotted points are shown in Exhibit 3.

Exhibit 6
Two Measures of Central Location in San Francisco

Notes: 3. Gazetteer. 4. Census 1982. Plotted points are shown in Exhibit 3.
Researchers who need access to a measure of city center locations have at least five databases at their disposal. Depending on the nature of the researcher's study, one may be more appropriate than the others; however, in some situations, the choice between them would likely not matter much, given the documented similarities across the measures. I conclude with the following suggestions for using these measures:

- The 1982 Census points are probably the best measures of the CBD concept; however, Fee and Hartley (2013) produced points for only 268 cities, and even the original survey delineated CBDs for only 455 cities. This source alone will not cover all principal cities in contemporary MSA samples.

- If the 1982 Census does not include the city being studied, use the City Hall database; these points come closest on average to the 1982 Census points in this analysis. These points are available for the principal cities of most of the currently defined MSAs.

- If a city is not available in either the 1982 Census or City Hall databases, the location of the city hall could be geocoded using Google Earth or ArcGIS.
If time constraints prohibit geocoding City Hall points, the Google Earth or ArcGIS geocoded city points could be used; the ArcGIS points were closer to the 1982 Census points in this analysis, suggesting that they are superior to the Google Earth measures for CBD purposes.

The Gazetteer should be used only in cases in which large measurement errors are not a problem, such as approximating airline travel distances between cities, or when they are used to measure the center of small census geographies, such as tracts in metropolitan areas, as in Holian and Kahn (2015). The Gazetteer measure is also the most readily accessible and comprehensive in geographic scope of the five measures.

Time permitting, researchers can always complete their tasks using multiple measures and then compare and contrast the results as a sensitivity analysis. I hope that continued research on city center location provides researchers with more and better data, and to that end I share the data and analysis file I used in writing this article.

Acknowledgments

The author thanks Rich Taketa and an anonymous reviewer, who offered helpful comments during the drafting of this manuscript. Any errors are the author’s.

Author

Matthew J. Holian is a professor of economics at San Jose State University and can be reached by email at matthew.holian@sjsu.edu.

References


