Assessing the Equity of Transit Supply Distribution in Metropolitan Areas Using Lorenz Curves and Gini Coefficients

Kelly L. Bertolaccini
kelly.bertolaccini@uconn.edu

Recommended Citation
http://digitalcommons.uconn.edu/gs_theses/483
Assessing the Equity of Transit Supply Distribution in Metropolitan Areas Using Lorenz Curves and Gini Coefficients

Kelly Layne Bertolaccini

B.S., University of Connecticut, 2010
B.A., University of Connecticut, 2010

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
At the
University of Connecticut
2013
Masters of Science Thesis

Assessing the Equity of Transit Supply Distribution in Metropolitan Areas Using Lorenz Curves and Gini Coefficients

Presented by
Kelly Layne Bertolaccini, B.S., B.A

Major Advisor
Nicholas E. Lownes

Associate Advisor
Carol Atkinson-Palombo

Associate Advisor
Karthik Konduri

University of Connecticut
2013
ABSTRACT

The equitable distribution of transit services is a major concern of transportation planners and policymakers worldwide. In the US, planners are required by law to consider equity concerns when investing in new transportation infrastructure and services. However, equity can be difficult to assess in a consistent, objective, and quantitative way. Australian researchers recently developed a single, system wide measure which reflects the horizontal (or spatial) equity of transit service distribution in a metropolitan area. This measure, a variation of the Gini coefficient, specifically measures how well transit supply meets transit demand. While using a single measure to assess the equity of a transit system is very appealing, researchers must be very careful when implementing Gini coefficients for comparative purposes. This research investigates the effect of using different scales, levels of data resolution, and various demand measures when calculating Gini scores for interregional comparisons. Gini coefficients are calculated for six urban transit systems at two scales (metropolitan statistical area and transit service area) and two levels of resolution (census tract and block group) using two different demand measures (population and population plus employment). The results suggest that calculating Gini coefficients at different scales can lead to drastically different comparative results, while the different levels of resolution and demand measures had very little impact on interregional comparisons. This research also explores the possibility of using Gini coefficients to assess vertical equity by using poverty data to estimate demand.
INTRODUCTION

In recent decades, transportation planners and policymakers have become increasingly interested in equity and its application in the development of transportation systems. A 2009 study led by Harvard economist Edward Glaeser found that residents of more unequal cities experience higher crime rates, a higher likelihood of political unrest and uprisings, and are more likely to report being unhappy (1). The social exclusion that can be caused by inequitable investments in transportation systems has become an international concern. Researchers, such as Delbosc and Currie (2), study the relationship between public transportation and social exclusion.

In the United States, planners and engineers are required by executive order to address environmental justice whenever planning or implementing a new project (3). Ideally, the benefits and burdens of transportation projects should be equally distributed among all populations, ensuring low income and minority communities are not treated unfairly (4). Any organization receiving federal funding, which includes most public transportation service providers, are also required to meet the equity requirements outlined in Title VI of the Civil Rights Act (5). Title VI of the Civil Rights Act states:

*No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.*

The Federal Transit Administration (FTA) requires all transit service providers to conduct equity analyses to ensure compliance with Title VI. Guidelines for these equity analyses can be found in the FTA Title VI circular (6).
Most planners interpret the executive order and Title VI legislation as a call to consider horizontal, or spatial, equity in the planning process. A horizontal equity perspective emphasizes the importance of treating people in equal circumstances equally. For example, a planner assuming a horizontal equity philosophy would suggest that all people living within a community deserve equal access to public transportation. When selecting projects, this planner would try to distribute projects evenly throughout the community without any group receiving a disproportionate amount of the burdens or benefits of transit projects. One of the major problems with the horizontal equity approach is that it fails to address or even consider existing inequalities. An alternative to the horizontal equity perspective is vertical equity which suggests that people in unequal circumstances should be treated unequally. A planner following a vertical equity philosophy would believe that disadvantaged populations, such as lower income families or ethnic minorities should receive priority consideration in public transportation projects. Unfortunately, these two types of equity often come into conflict. If disadvantaged groups are being prioritized, then everyone is not being treated equally.

In her paper on incorporating environmental justice (EJ) into the traffic assignment step of the traditional transportation planning process, Jennifer Duthie discusses this issue (7). The paper attempts to incorporate EJ measures into the user equilibrium-based discrete network design problem (UE-DNDP). In this research, Duthie tests eight possible objective functions for the program formulation, with four taking a vertical equity perspective, giving greater weight to projects serving protected populations, and the other four taking a horizontal equity perspective, treating all populations equally. She found that the vertical equity and horizontal equity objectives often came in conflict with one another. The objective function that was specifically designed to equally distribute system benefits among all populations, regardless of protected
status, actually provided the least amount of improvements for all population groups. In other words, rather than redistributing the extra benefits that would have been given to protected populations throughout the community, the objective function just cuts the extra projects. Most people would agree this is not an ideal solution. While Duthie’s findings are specific to her formulation of the highway network design problem, the implications of designing for horizontal equity over vertical equity extend into other areas, including transit planning. Public transportation adds another complication to the horizontal v. vertical equity problem: transit planners need to consider ridership and not all communities or demographics are equally likely to use transit. A TCRP report found that lower income households are more likely to choose transit over auto than higher income households (8). Though this research will primarily focus on using the Gini coefficient as a measure of horizontal equity, it is important to recognize the limitations of taking that perspective. This research will explore some possibilities for expanding the Gini coefficient to measure vertical equity but these methods require further research before application.

Recently, Delbosc and Currie (9) suggested a single measure to describe the distribution of transit supply among the transit demand of a metropolitan region. This measure is a Gini coefficient adapted to reflect the distribution of transit supply and can be considered a measure of horizontal equity. Traditionally, the Gini coefficient compared a population’s distribution of income, represented as a Lorenz curve, to a line representing perfect equality. The Lorenz curve plots the cumulative proportion of the population, ordered from lowest to highest income, against the cumulative proportion of income earned (10). When using the Gini to approximate the distribution of transit supply, Lorenz curves are plotted as the cumulative proportion of transit demand, ordered by demand density, against the cumulative proportion of transit supply. A
perfectly equal distribution of supply would result in a Gini coefficient of 0 while a perfectly unequal distribution of income would result in a coefficient of 1.

There are some caveats to using the Gini coefficient. A region’s Gini coefficient cannot be interpreted on its own; rather it must be used within a comparative framework. There are limitations to the types of comparisons that can be made with the Gini. Gini is very sensitive to issues of scale, resolution and sample size, making it difficult to make meaningful interregional comparisons. This sensitivity to subjectively or arbitrarily defined geographic regions is commonly referred to as the modifiable areal unit problem (MAUP) (11). The MAUP is often divided into two sub problems: the scale effect and the zonal effect (12). The scale effect refers to the variations in data and results due to the aggregation of data at different geographic scales. For example, a Gini coefficient calculated using block group level population data will differ from a Gini coefficient calculated using the same population data aggregated at the census tract level. The zonal effect refers to the variations in data and results due to the arbitrary nature of selected boundaries (13). For example, the calculated Gini coefficient will change depending on how block group boundaries are drawn through a population. In this paper, the MAUP will be explored by calculating Gini coefficients at two different resolutions: block groups and census tracts. The transit Gini coefficient will also be affected by the area considered in the analysis which will be referred to in this paper as the scale.

This paper will apply a modified version of the transit Gini coefficient developed by Delbosc and Currie to six metropolitan areas within the United States. One of the purposes of this particular study is to examine the effects of calculating Gini at different scales and resolutions using different demand measures. This will provide valuable insight to planners and researchers attempting to use this coefficient for inter-city comparison. The following section
will provide a detailed description of the methodology used to calculate the Gini coefficient, including the calculation of the transit supply index, the estimation of transit demand, and the development of Lorenz curves. Next, the results of this analysis will be reported, interpreted and discussed. The final section will report overall conclusions and suggest topics for further research.

METHODOLOGY

Calculating a Gini coefficient to reflect the distribution of transit supply requires developing a Lorenz curve which plots the cumulative proportion of transit supply against the cumulative proportion of transit demand. Transit supply and demand are estimated at two different resolutions, block groups and census tracts. This study uses a transit supply index to quantify supply and a combination of population and employment to estimate the demand of each areal unit. The following sections will describe the supply and demand calculations in greater detail. The results of these calculations are then used to develop the Lorenz curve and calculate the Gini which is discussed in the final portion of the Methodology section.

One of the purposes of this study is to evaluate the effects of geographic scale and resolution on a city’s calculated Gini coefficient. In this study, Gini coefficients are calculated at two different scales (Metropolitan Statistical Area (MSA) and Transit Service Area (TSA)) and at two different levels of resolution (census tract and block group). The TSA was defined as the set of block groups or census tracts within the MSA that received any amount of transit service. Additionally, two different methods are used to estimate transit demand (population and population plus employment), resulting in eight different Gini coefficient calculations for each of the selected cities.
The following six MSAs were selected for inclusion in this study:

- Albuquerque, NM
- Austin, TX
- Cleveland, OH
- Denver-Boulder, CO
- Madison, WI
- Wilmington, NC

Though the Denver-Boulder region is classified by the US Census Bureau as two separate MSAs, they are considered together in this study because they share a single transit system. These MSAs and transit systems were specifically selected to generate a diverse set of case studies, representing cities and systems of different sizes and densities in various regions of the United States. To be considered for selection in this study, a city needed to be located entirely within one state and provide recently updated transit system data in the commonly used General Transit Feed Specification (GTFS) format. GTFS is the format used by Google Transit and many independent app developers (14). GTFS allows for common data collection periods, consistent interpretation and simpler processing of data.

Supply

This paper uses a transit supply index largely based on previous efforts in Australia (9,15). The transit supply index accounts for station coverage and vehicle frequency, but does not consider how connected a particular stop is to the rest of the network or how transit supply varies throughout the day. While this measure of supply is far from comprehensive, the ease of
calculation makes it a practical choice for practitioners. Equation 1 below shows how the transit supply index was calculated for each geographic unit (either block group or census tract).

\[
TSI_{GU} = \sum_{R} \left( \frac{Coverage Area_{R, GU}}{Total Area_{GU}} \times f_{R, GU} \right) \tag{1}
\]

Where \( f_{R, GU} \) is the mean weekly vehicle arrivals for stations along route R in geographic unit GU. The coverage area of route R in geographic unit GU refers to the total area within the geographic unit that is within 0.25 miles of a bus station or 0.5 miles of a rail station. An example included later in this section will illustrate how the transit service index is calculated for a particular geographic unit. The major difference between this index and the one developed by Currie is that rather than developing separate scores for each mode, this index develops separate scores for each route. This alteration helped to address the problem of differing weekly frequencies between stations located within the same geographic unit. However, this did not fully solve the problem and many stops along the same route had differing frequencies. This problem was addressed by finding the mean weekly vehicle arrivals for stations along a route within a specific geographic region.

Calculating the Transit Supply Index requires knowing the location of each transit station and the number of vehicles that stop there each week. GTFS data was used to generate a single table including the latitude, longitude, and the number of vehicles per week per route of each station. When the data was brought into the GIS environment to generate a point shapefile of station locations, a separate point was generated for each route using a specific stop location. For example, if one station serviced three bus routes, then three separate points would be generated at the location, one for each station-route combination. Generating a separate point for each station-route combination is important for the TSI calculations. In addition to the transit system
data from the GTFS data, calculating the TSI also requires block group/census tract shapefiles for the metropolitan area.

After generating the station-route shapefile with weekly frequencies, ArcGIS was used to complete the following steps and calculate the TSI of each areal unit:

1. Generate 0.25 mi and 0.50 mi buffers around bus and rail stations, respectively, dissolving by route and weekly frequency
2. Perform a union between the buffer layer and the areal unit (block group/census tract) layer. Dissolve by route and areal unit finding the mean weekly frequency for each route through a particular areal unit.
3. Find the percentage of each areal unit covered by a route’s dissolved buffer and multiply by the mean weekly frequency. This calculates the supply score of one route through an areal unit.
4. Sum all supply scores within an areal unit to find the total TSI of each block group/census tract.

The example below illustrates how to calculate the TSI of a sample block group /

*Example: Calculating Transit Supply Index*

A sample block group is shown in Figure 1 below. Two bus routes (A and B) and one rail line serve this block group. Figure 1 includes the buffers generated in Step 1 of the procedure.
Figure 1: Sample Block Group

Steps 2 and 3 of the outlined procedure will give the results shown in Figure 2 below.

Figure 2: Calculating Route Transit Supply Indices
To find the total TSI of the block group, sum the supply scores found for each of the routes. This block group has a TSI of 82.25.

**Demand**

Demand was estimate two ways: population and population plus employment. All population data, for both the block group and census tract levels, came from the 2010 US Census. Employment data came from the 2010 American Community Survey (ACS) five year estimates for place of work. However, employment data was only available at the census tract level. For block group calculations, employment was assumed to be evenly split amongst the block groups. All of the demographic data was accessed through the US Census Bureau’s American Factfinder website.

**Lorenz Curve and Gini Calculations**

In this study, the Gini coefficient quantifies how well transit supply is distributed among the transit demand by comparing the actual distribution of transit supply, represented as a Lorenz curve, to a line representing a perfectly even distribution, the line of equality. The Lorenz curve plots the cumulative proportion of transit demand against the cumulative proportion of supply. A point on the Lorenz curve can be read as X percent of the demand receives Y percent of the supply. Figure 3 shows an example of a Lorenz curve plotted against a line of equality.
Equation 2 shown below is used to calculate the Gini coefficient from the Lorenz curve and line of equality.

\[
Gini = \frac{A_{equal} - A_{Lorenz}}{A_{equal}} \quad (2)
\]

Where \( A_{Lorenz} \) is the area underneath the Lorenz curve and \( A_{equal} \) is the area underneath the line of equality. Because the axes are scaled from 0 to 1, \( A_{equal} \) will always equal 0.5. A perfectly even distribution of supply would result in a Gini coefficient of 0 while a perfectly unequal distribution of supply would result in a coefficient of 1. Note that a perfectly even distribution of transit supply does not imply that demand for transit service is being perfectly met. A transit system can receive a “perfect” score of zero regardless of how well the system is able to meet the total demand.
As previously mentioned, demand was estimated in two different ways (population and population plus employment). However, for ease of comprehension, the following portion of text, describing the creation of Lorenz curves, will simply refer to population. The same procedure is used to generate Lorenz curves for demand measured as population plus employment. The x-axis of the Lorenz curve plot displays the cumulative proportion of the population, ranking geographic units in order of increasing demand. Ranking the geographic units presented a challenge. The geographic units with the highest total population were often the large, low density units located outside of the urban core which one would reasonably expect to receive less transit service than the smaller, high density units inside the city core. However, using population density rather than total population as the measure of transit demand makes the Lorenz curve impossible to interpret. To resolve this issue, the Lorenz curves generated in this study used total population as the measure of demand but ranked the geographic units in order of population density. After generating the Lorenz curves, equation 2 above was used to calculate the Gini coefficients.

**DISCUSSION OF RESULTS**

This analysis generated many interesting maps and graphs, only some of which will are included in this report. The first portion of the discussion summarizes the distribution of supply and demand measures over the selected metropolitan areas. Each metropolitan region is described using four maps which show: (a) population density, (b) employment density, (c) transit supply at the MSA scale, and (d) transit supply at the TSA scale. With the exception of the Denver-Boulder MSAs, population and employment density maps are shown at the MSA
scale. The next portion of this section will present overall results and explore the effects of using the different Gini calculation methods to make comparisons between cities. Then it will delve into some of results from individual cities to explain how planners can utilize the Lorenz curves and Gini coefficients in the planning process. The final portion of this section discuss the possibility of incorporating poverty, income, and other demographic data into the demand measure to transform the Gini coefficient into a measure of vertical equity.

The following pages show detailed maps of the distribution of supply and demand measures in the selected case study metropolitan areas. While the information displayed in these maps has many implications for each of the individual metropolitan regions and their respective transit services, this discussion will focus on some of the important trends observed across regions. As expected, block group’s and tract’s with high population densities also tended to have high employment densities. However, it is important to recognize that the similarities in the spatial distribution of population and employment densities do not imply that people are actually residing near their place of employment. People working in areas of high employment density often commute from lower density residential areas outside of the urban core. At the same time, people living in densely populated urban areas may be commuting to service jobs in lower density suburban areas. Furthermore, areas with high levels of transit service coincide with the
areas of high population and employment density. This will result in relatively low Gini coefficients for the metropolitan regions, indicating a highly equitable system, regardless of the transit service’s actual ability to connect people’s homes and places of work. The maps also reveal variation in the relative size of a transit service area to the size of the metropolitan region it serves. The implications of this variation will be discussed later in this section.

After calculating TSI scores and gathering population and employment data, eight Gini scores were calculated for each of the selected transit systems. Table 1 below shows all eight of the Gini coefficients generated for each city.

Table 1: Gini Coefficient Results

<table>
<thead>
<tr>
<th>City</th>
<th>Scale</th>
<th>Resolution</th>
<th>Gini (Population)</th>
<th>Gini (Population+Employment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albuquerque</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.1451 (1)</td>
<td>0.1706 (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.1448 (1)</td>
<td>0.1267 (1)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.3441 (2)</td>
<td>0.3569 (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.3371 (1)</td>
<td>0.3223 (1)</td>
</tr>
<tr>
<td>Austin</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.2619 (3)</td>
<td>0.2848 (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.3388 (4)</td>
<td>0.3348 (4)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.5316 (4)</td>
<td>0.5432 (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.5762 (4)</td>
<td>0.5718 (4)</td>
</tr>
<tr>
<td>Cleveland</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.2412 (2)</td>
<td>0.2353 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.2800 (2)</td>
<td>0.2519 (2)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.4625 (3)</td>
<td>0.4545 (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.4733 (3)</td>
<td>0.4501 (3)</td>
</tr>
<tr>
<td>Denver-Boulder</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.2796 (4)</td>
<td>0.3090 (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.3079 (3)</td>
<td>0.3087 (3)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.3422 (1)</td>
<td>0.3684 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.3567 (2)</td>
<td>0.3564 (2)</td>
</tr>
<tr>
<td>Madison</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.5080 (6)</td>
<td>0.5230 (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.5603 (6)</td>
<td>0.5545 (6)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.7124 (6)</td>
<td>0.7198 (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.7275 (6)</td>
<td>0.7231 (6)</td>
</tr>
<tr>
<td>Wilmington</td>
<td>Service Area</td>
<td>Block Group</td>
<td>0.4214 (5)</td>
<td>0.4592 (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.4911 (5)</td>
<td>0.4876 (5)</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>Block Group</td>
<td>0.6833 (5)</td>
<td>0.7010 (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tract</td>
<td>0.7088 (5)</td>
<td>0.7058 (5)</td>
</tr>
</tbody>
</table>
The numbers shown in parentheses indicate the rank of the specified city when the Gini coefficient is calculated using the indicated method. The cities are ranked from the most evenly distributed transit supply (low Gini coefficient) to the least evenly distributed transit supply (high Gini coefficient). Because Gini coefficients tend to be used in comparison, the changes in rankings due to different calculation methods are particularly important. Figures 4, 5, and 6 show how the use of different scale, resolution, and demand measures, respectively, affect the calculated Gini coefficients.

**Figure 4: Comparison of Gini coefficients at different scales calculated using block group level resolution with population as demand.**
Figure 5: Comparison of Gini coefficients at different resolutions calculated at the service area scale with population as demand.

Figure 6: Comparison of Gini coefficients using different demand measures at MSA scale and tract level resolution.
Using different scales to calculate Gini had the greatest impact on the both the overall values of the coefficients and the rankings of the city coefficients, as shown in Figure 4. This is not surprising considering the variations in the relative size of the transit service area to the metropolitan statistical area. For example, the Wilmington TSA covers only 9% of the MSA’s area and 44% of the MSA’s total population. It is therefore unsurprising that calculating the Gini coefficient at these two different scales led to drastically different results. In contrast, the Denver-Boulder TSA covers 18% of the MSA’s area, 87% of the MSA’s population, and experienced a much smaller change in Gini calculated at the different scales. This contrast points to the importance of carefully selecting the scale of analysis. MSAs are defined by the US Office of Management and Budget as including “at least one urbanized area of 50,000 or more in population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.” (16) This vague definition may cause the MSAs to have very different characteristics in terms of population density and distribution, calling into question the validity of using the MSA as a scale of comparison. Researchers and planners intending to use the Gini as a comparative tool should be very careful and specific when selecting an appropriate scale of analysis.

The different levels of resolution and demand measures appear to have a much less significant impact on the Gini coefficient. Calculating Gini at the block group level consistently resulted in a slightly lower score than calculating Gini at the census tract level. Including employment in the demand measure resulted in slightly lower Gini coefficients at tract level resolution as expected. The opposite was true of including employment at block group level resolution, but given the low resolution of employment data, nothing should be read into this counterintuitive result. As long as researchers use a single level of resolution and select a
demand measure consistent with that level of resolution, their choices should not have a significant impact on city comparisons.

While Gini coefficients lend themselves to making comparisons between cities, the Lorenz curves themselves can tell planners a lot about the current distribution of transit supply within a select city. Figures 7(a) and (b) below show some of the Lorenz curves generated by the Denver-Boulder and Madison case studies.

(a)
Figure 7: Lorenz Curves for (a) Denver-Boulder (b) Madison. Curves calculated at block group level resolution with population as demand.

The Denver-Boulder transit system is an example of a system that was designed to serve an entire metropolitan region. The vast majority of the population, including those living outside the urban core, live within the TSA, suggesting that the planners intended the system to be used by both commuters and city residents. In other words, transit supply was intentionally distributed to accommodate suburban demand. This is reflected in the smooth, flat Lorenz curve and relatively low MSA scale Gini coefficient. In contrast, the Madison transit system focuses on serving those living and working in the urban core rather than longer distance commuters. Though the block groups located outside of the urban core have relatively low population densities, these block groups also cover much larger areas than those in the core and therefore contain a significant proportion of the MSA’s population. Madison’s MSA Lorenz curve does not leave the x axis until reaching approximately 40% of the demand, or population. Because a
significant proportion of the MSA receives no transit supply, Madison’s MSA scale Gini coefficient is relatively high. If planners decide that their city’s transit system should be focused on people traveling within the urban core rather than long distance commuters, they should expect high MSA scale Gini coefficients. Planners must consider their city’s transit goals when selecting a target Gini coefficient and scale of analysis.

The Gini coefficient could potentially be incorporated into the project selection process as well. Using Gini to compare proposed transit projects within a given MSA avoids many of the complications of inter-city comparisons. Though a thorough analysis of this possibility was not part of this study, a small experiment was conducted on Austin’s transit system to see whether making changes to one route in a large system would have a noticeable impact on the Gini coefficient. In this experiment, the weekly vehicle frequency along Austin’s commuter bus route No. 983 was doubled. This resulted in an approximate 0.02 increase in Gini coefficients at the TSA scale. There was no noticeable change at the MSA scale. Increasing the vehicular frequency along this existing route actually resulted in a slightly less equal distribution of transit supply among the population. Further research is required to determine how adding new stops and routes effects Gini and how to use the Gini coefficient to pinpoint the areas in greatest need of increased transit services.

To transform the Gini coefficient from a measure of horizontal equity to a measure of vertical equity, demand estimates need to account for some of the demographic differences among the population. For this experiment, poverty density persons living below the poverty threshold/sqmi replaced population and employment density as the measure of demand. The maps on the following pages compare the distribution of poverty density and transit supply in the selected cities’ transit service areas. Poverty data was only available at tract-level resolution.
As expected, poverty density appears to be distributed similarly to population density. However, using poverty density to estimate demand yielded different Gini coefficients and altered the rankings between the states, as shown in Table 2 below. This table also includes the coefficient of determination ($R^2$) of the relationship between poverty density and population density to assist in the interpretation of these results.

### Table 2: Gini Coefficients (Poverty Density and Population Density)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Albuquerque</td>
<td>Service Area</td>
<td>0.5704</td>
<td>0.1491</td>
<td>0.1448</td>
<td>0.5333</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.4595</td>
<td>0.0914</td>
<td>0.3371</td>
<td>0.5949</td>
</tr>
<tr>
<td>Austin</td>
<td>Service Area</td>
<td>0.4947</td>
<td>0.0682</td>
<td>0.3388</td>
<td>0.6516</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.3522</td>
<td>0.2959</td>
<td>0.5762</td>
<td>0.6252</td>
</tr>
<tr>
<td>Cleveland</td>
<td>Service Area</td>
<td>0.4122</td>
<td>0.1758</td>
<td>0.2800</td>
<td>0.3938</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.4926</td>
<td>0.0231</td>
<td>0.4733</td>
<td>0.4162</td>
</tr>
<tr>
<td>Denver-Boulder</td>
<td>Service Area</td>
<td>0.5311</td>
<td>0.0663</td>
<td>0.3079</td>
<td>0.5992</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.5173</td>
<td>0.0429</td>
<td>0.3567</td>
<td>0.6042</td>
</tr>
<tr>
<td>Madison</td>
<td>Service Area</td>
<td>0.4809</td>
<td>0.0794</td>
<td>0.5603</td>
<td>0.8763</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.3780</td>
<td>0.2440</td>
<td>0.7275</td>
<td>0.8613</td>
</tr>
<tr>
<td>Wilmington</td>
<td>Service Area</td>
<td>0.4072</td>
<td>0.1857</td>
<td>0.4911</td>
<td>0.6774</td>
</tr>
<tr>
<td></td>
<td>Metro Area</td>
<td>0.2451</td>
<td>0.5099</td>
<td>0.7088</td>
<td>0.6842</td>
</tr>
</tbody>
</table>

One interesting difference these results and the previously results is that the area under the Lorenz curves was greater than 0.5000 in the Albuquerque TSA and the Denver-Boulder MSA and TSA. This means that tracts with low poverty densities received larger proportions of transit service than tracts with higher poverty densities. Scale of analysis does not appear to have any consistent effect on the Gini coefficients and, in fact, the impact of scale also varied between demand measures. For example, Cleveland’s population-based Gini coefficient increased when the scale shifted from TSA to MSA, while its poverty-based Gini did the opposite. Upon a closer
examination of the distribution of poverty and population in the Cleveland metropolitan area, this
variation in Gini score makes sense. People living below the poverty line are much more heavily
concentrated in the urban core, with high levels of transit service, than the general population.
The $R^2$ value of poverty density and population density does not seem to provide any additional
insight into the changes in Gini coefficients with scale and demand measure.

Unfortunately, there are several very serious problems with using poverty density as a
proxy for demand. The first is with the way poverty is defined. A family’s poverty status is
determined by whether or not the family’s total income (before taxes) falls below the national
thresholds set by Office of Management and Budget and the US Census Bureau. These
thresholds vary by the size of the family and the age of its members and are adjusted annually
using the Consumer Price Index for All Urban Consumers (17). The use of thresholds turns
poverty, which seems best understood as a continuum, into a discrete variable. This is
problematic because there is no real difference in the level of poverty experienced by a family
living one dollar below the poverty line and another living one dollar above it. Another major
concern is that highest level of resolution for poverty data is the census tract. It is unrealistic to
assume poverty is uniformly distributed throughout an area as large a census tract.

However, one of the most significant concerns with poverty density as a demand measure
is that, similar to population density, it fails to account for whether or not the transit service
actually connects potential users’ residences and places of work. The Census Bureau does not
publish poverty status or income data based on place of work. The Census Transportation
Planning Package (CTPP) does provide this information, as well as actual origin-destination pair
data, but has not been updated since 1999 so it was incompatible with the transit supply data
used in throughout this study. Without incorporating origin-destination data into the demand
measure, the Gini Coefficient is unlikely to reflect how well a transit service is actually meeting the needs of the population. The new CTPP, based on data from the 2010 Decennial Census, is set to be released in the spring of 2013. Future research should consider incorporating origin destination data into the demand estimates.

CONCLUSIONS

Lorenz curves and Gini coefficients can be used to assess how evenly transit supply is distributed among demand. Researchers and planners applying the Gini coefficient must make decisions on the scale of analysis, level of data resolution, and demand measurements. These decisions affect not only the actual transit Gini coefficient calculated for a particular region but can also affect interregional comparisons. This analysis calculated Gini coefficients for two scales (MSA and TSA) at two levels of resolution (census tract and block group) using two different demand measures (population and population plus employment).

The results showed that shifting between levels of resolution and demand measures had a relatively consistent impact on the calculated Gini coefficients. Therefore as long as researchers are consistent in applying the same demand measurements and calculating Gini at the same level of resolution, interregional comparisons between Gini coefficients should be valid. However, researchers must be much more careful in selecting a scale of analysis which can have a drastic effect on interregional comparisons. Using subjectively defined geographic boundaries, such as Metropolitan Statistical Areas, as a scale of analysis can have an inconsistent effect on Gini calculations. If a transit system is designed for commuters and serves a large proportion of the MSA’s population, then the difference between Gini coefficients calculated at the TSA and MSA scales will be relatively small. However, if the transit system is designed solely for use
within the urban core and does not serve those living outside of the city, the difference between Gini coefficients calculated at the TSA and MSA scales will be much larger. Researchers must be very careful to select an appropriate scale for their analyses.

Future research should investigate the possibility of using the transit Gini coefficient as a planning tool. Gini coefficients could potentially be used to compare how proposed projects will affect the distribution of transit supply in the region. Future research should also consider using more sophisticated measures of demand. If a demand measure incorporating demographic variables such as household income and car ownership replaced the current demand measure of total population and employment, the Gini coefficient could then be used to assess vertical equity and address environmental justice concerns.
REFERENCES


3. Executive Order 12898, 1994, President William Clinton.


   <http://www.fta.dot.gov/documents/FTA_Title_VI_FINAL.pdf>


10. Gini, C.W. Variability and Mutability, Contribution to *the Study of Statistical*


