Chapter 1
I Dream of Gini: Measures of Population Concentration and Their Application to US Population Distribution

Peter A. Rogerson

Abstract The unevenness of population across space may be captured with various measures of inequality; the Lorenz curve, the Gini coefficient, and the Hoover index constitute prime examples of such measures. In this chapter, I first review the history of these measures and then provide a selective review of their use in examining population concentration and deconcentration. Next, I show how the Gini coefficient may be disaggregated to show how population concentration varies within different ranges of population densities. The Gini coefficient is written as a weighted sum of the Hoover indexes for each population density category, where the weights are the proportion of total area in that density category. This disaggregation is applied to the US population for the period 2000–2015. Results show that population deconcentration is occurring among the subset of counties that have high population density and concentration is occurring among counties that have medium population density.

Keywords Gini coefficient · Hoover index · Population deconcentration

I am pleased to be able to contribute to this volume in honor of David Plane's contributions to the fields of geography and regional science. Dave and I first met in 1980 at the US Census Bureau, where we were doctoral students working on a project jointly funded by the American Statistical Association, the National Science Foundation, and the Census Bureau. Since that time, we have enjoyed both a career-long collaboration that has been productive and rewarding and a lifelong friendship that has been enjoyable and genuine. Over the course of our careers, we worked together on many topics at the intersection of demography and geography. Some of our early direction and inspiration came from Andrew Isserman, the leader of our project at the Bureau. Following Andy's passing in 2010, Dave and I coauthored a

P. A. Rogerson (✉)
Departments of Geography and Biostatistics, Wilkeson Hall, University at Buffalo, Buffalo, NY, USA
e-mail: rogerson@buffalo.edu

paper on the Hoover index of population concentration for a special volume of the *International Regional Science Review* that was put together in memory of Andy. With this chapter, I am delighted to return to this topic. A special word about the title—over the years, Dave and I shared both hotel rooms at conferences and tents in campgrounds while on bike trips. In addition to the many stimulating conversations, Dave learned that I often have vivid dreams—these ranged from falling out of both my bed and then our 35th story window at the Bonaventure Hotel in Los Angeles to shouting away the intruders of my dreams that lurked outside of our tent. I think, then, it is only fitting to include an allusion to this in the title of the chapter.

1.1 Introduction

Substantial change has taken place in the distribution of population across counties over time in the United States. On a long time scale, the population has of course become urbanized, rising from 5.1% urban at the time of the first decennial census in 1790 to 80.7% urban in 2010. An increase has occurred in every decade with the exception of 1810–1820, when the urban population went from 7.3% to 7.2% (US Bureau of the Census 2012).

With respect to more recent change, a plethora of studies emerge whenever new census data are released. A Pew research report notes that urban counties have grown at about the nationwide rate of 13% during the period 2000–2016, while suburban and small urban counties have grown more rapidly (16%) (Parker et al. 2018). Simultaneously, rural counties have grown more slowly, rising just 3% over the period.

Similarly, Frey (2018), in a report from the Metropolitan Policy Program of the Brookings Institution, notes that urban core counties grew more slowly (at an annual rate of approximately 0.5%) than did mature and emerging suburban counties and exurban counties (annual rate of a bit over 1%) during the period 2000–2017. Toward the end of that period, small metropolitan areas began to grow at the expense of large metropolitan areas.

A widely recognized confounding factor in these studies is the definition of terms such as rural, suburban, urban, and exurban. Definitions vary from study to study, and the census definition itself has varied over time. Hall, Kaufman, and Ricketts (2006) provide an overview of many of the available definitions.

1.2 Some Measures of Concentration: With Historical Notes

One of the earliest measures of concentration was described by Lorenz in 1905; the well-known Lorenz curve provides a convenient way to visualize inequality. It is most commonly used for the depiction of income inequality. In that context, after

arranging individuals (or groups of individuals) in terms of increasing income, cumulative population is plotted on the horizontal axis against cumulative income on the vertical axis. When there is little inequality, the plot will lie close to the 45° line. As inequality increases, the Lorenz curve bows further outward and away from (and below) the 45° line.

Shortly after, Gini (1912, 1914) introduced a numerical measure of inequality that is directly related to the Lorenz curve. In particular, the Gini coefficient is equal to the fraction of the area lying below the 45° line that is between the 45° line and the Lorenz curve. The Gini coefficient is also equal to half of the relative mean absolute difference between all pairs of incomes:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \sum |x_i - x_j|}{2n^2\bar{x}}$$

The measure is variously known as the Gini coefficient (1,170,000), Gini index (871,000), and the Gini ratio (141,000), where the numbers in parentheses indicate the number of "hits" in Google (recognizing of course that this is not a perfect indicator of popularity or use).

As noted by Ceriani and Verme (2012), Gini (1912) first described his index in terms of mean differences, in a publication in Italian that has never been translated into English. In that publication, Gini discusses 13 different measures of inequality, of which the Gini coefficient is one. He draws an explicit connection to the Lorenz curve in his 1914 publication—this was eventually published in English, but not until 2005 (Gini 2005). Ceriani and Verme suggest that Gini's work became more widely known and disseminated following a note by Gini (1921) in English in *Economic Journal*, where he brought readers to the attention of the work of several Italian researchers. The note focused on measures of inequality and was written in response to an article on the topic in that journal by Dalton (1920).

The Hoover index of concentration, as it is known to geographers and demographers, is equal to the maximum vertical distance between the Lorenz curve and the 45° line. It is also equal to half of the relative mean absolute deviation (from the mean):

$$H = \frac{\sum |x_i - \bar{x}|}{2n\bar{x}}$$

In the context of income, it is interpreted as the percentage of total income that would have to be redistributed from the rich (defined as those to the right of the vertical line) to the poor (defined as those to the left of the vertical line), to equalize incomes. In the context of population, it is the percentage of population that would have to move from high-density places to low-density places to equalize population density across all spatial units. The Hoover index (13,000) is known variously as the Pietra index (3000), the Robin Hood index (15,000), and the Schutz index (47,000), where the number in parentheses is again the rounded number of "hits" in a Google search.

Hoover's 1936 paper is often cited as the one where he introduced what is now known as Hoover index of concentration. However, in that paper he used the Gini index, and there is no mention of his index of concentration. He cited Gini as an important contributor to measures of inequality and thanked the Nobel Prize economist Wassily Leontief for pointing this out to him, but he does not appear to have noticed at that time that Gini was responsible for the development of the coefficient; instead, he credited Vinci (1934) for interpreting the coefficient.

Hoover introduced his index of concentration in his 1941 paper. Here he credits Florence and Wensley (1939) for developing the measure, citing their book in the context of the extent of localization of manufacturing industries. Florence (1953) himself notes that he and Wensley first reported their initial work on this measure in 1937, in *Economic Journal*. And interestingly, in this 1937 work, they cite the work of Hoover as one of several researchers working on *other* measures.

Hoover's index actually appears to have first been discussed by Pietra (1915; translated in Pietra 2014), who showed the relations between several measures of inequality/concentration. Giorgi (2014) notes that Pietra showed the "relations existing between the mean difference, the simple mean deviations from the arithmetic mean and the median, and their geometrical interpretation". Schutz (1951) was a graduate student at Berkeley when he published his explication of the index (apparently without knowledge of and certainly without attribution of the earlier work of Pietra, Florence and Wensley, and Hoover). Despite the existence of earlier formulations, the term "Schutz index" is used much more often than the other names for the measure, perhaps due to its introduction and use within the field of economics.

1.3 Selected Review of Studies of Population Concentration and Deconcentration in the United States

The following review is meant to be illustrative and representative of the types of studies that have focused on population concentration and its measurement in the United States; it is by no means meant to be a comprehensive review.

Duncan et al. (1961) used the Hoover index to show how the US population became increasingly concentrated during the first half of the twentieth century. Lichter (1985) also used the Hoover index, focusing upon the latter half of the century, and the differing degrees of concentration by race.

Plane and Mulligan (1997) argue for the use of the Gini coefficient in population and migration research. They calculate and interpret several Gini coefficients in the context of the US migration system, concentrating on the measurement of the amount of spatial focusing that occurs—either within the entire system or within the sets of inflows and outflows. Rogers and Sweeney (1998) and Rogers and Raymer (1998) apply these measures and the coefficient of variation in their own analyses of US population redistribution.

Lichter and Johnson (2006) examine the spatial patterns of concentration and deconcentration among the foreign-born population that occurred during the 1990s. They find that (a) the foreign-born are dispersing away from metropolitan, gateway cities (although they remain much more concentrated than the native-born) and (b) they are less segregated from other populations than they were in the past—so-called balkanization and isolation are not as acute as they once were.

Plane et al. (2005) examine population distribution in the context of the urban hierarchy. They place their work in historical context by noting that over the long-term net migration has been primarily up the urban hierarchy, leading to increasing population concentration. They use migration data for the period 1995–2000 to look at net flows up and down the hierarchy—the nature of such flows can differ substantially according to age and life course stages and events (such as college, military service, family formation, retirement, etc.).

By far, the majority of attention to population concentration and deconcentration in the United States has been devoted to the "rural renaissance" of the 1970s and the subsequent relative growth rates of rural, urban, and suburban areas. Vining and Strauss (1977) argued that the recently observed deconcentration was a "clean break" from past trends. At about the same time, McCarthy and Morrison (1977) also noted the increased net in-migration that was being experienced by rural, nonmetropolitan counties, signaling the beginning of a new or at least more complex demographic trend. They also found that retirement and recreation were increasingly important as drivers of this migration.

Gordon (1979) argued that the newfound reversal was perhaps due to growth that was extending from metro areas and spilling over into adjacent nonmetropolitan counties and that the "clean break" hypothesis deserved further scrutiny. He used the Hoover index to support this view using data from 18 countries. Bourne (1980) attributed the decreasing concentration to implicit US policies that favored the growth of exurbia (and the lack of policies that encouraged redevelopment within cities).

Work in this area developed rapidly, with Berry (1980) providing an early review of the counterurbanization trend and John Long (1981) producing a book describing the trends toward deconcentration at various spatial scales. Lichter and Fuguitt (1982) focused some of their attention on population distribution *within* nonmetropolitan areas, showing that there was deconcentration at that level as well. Like Morrison and McCarthy, they found that economic explanations for demographic change in nonmetropolitan areas were of decreasing importance.

The empirical work also led to the development of various theoretical frameworks within which the changes could be understood more broadly. Morrill (1979, 1980) and Geyer and Kontuly (1993) developed conceptual and theoretical frameworks for population and concentration; the latter authors, for example, use data from different countries to suggest that counterurbanization represents the final phase of the first cycle of demographic change in an urban system, and this is followed by a cycle where concentration is again dominant. Morrill (1980) argued against a "clean break" in the 1970s, suggesting that long-term agglomerative forces acted differentially across space and that older and denser places experience out-migration and deconcentration as a stage in the evolution of the geographic landscape.

Fuguitt (1985) was an early reporter of the reversal of the turnaround, reporting that there was a return to concentration in the metropolitan nonmetro system during the early 1980s. Fuguitt also reviewed the literature on the turnaround, emphasizing the point that the focus on trends in concentration and deconcentration had the beneficial effect of drawing more attention to other facets of migration research, including individual migration behavior, preferences, and the relation of migration to employment. Finally, he pointed out that the research generated by the topic was not only broad in scope but also large in its volume. At that time Fuguitt speculated that a return to concentration seemed unlikely. Cochrane and Vining (1988) also provided early evidence of the reversal of counterurbanization that occurred during the 1980s.

Then, during the 1990s, there was a return to the deconcentration witnessed in the 1970s. Long and Nucci (1997a) documented this return to deconcentration using the Hoover index, with counties as the spatial unit. In a second paper, Long and Nucci (1997b) documented this further, simultaneously (a) correcting an error in the original county-based series of Hoover indexes reported by Duncan et al. (1961), (b) extending the Duncan et al. time series both backward and forward in time, and (c) noting that there was also deconcentration at the county level between 1890 and 1910. Johnson and Cromartie (2006) provided corroborating evidence for the renewed deconcentration that took place during the 1990s by using data from the 2000 census.

Domina (2006) found that trends changed again during the late 1990s and the early years of the twenty-first century, with nonmetro areas once again experiencing net out-migration. Domina used data from the Current Population Survey to show that much of the net out-migration was attributable to those with higher levels of education and to suggest that economic factors now carry important explanatory power in understanding nonmetro migration.

Rogerson and Plane (2013) calculated annual Hoover indexes for the period 1990–2009 at a variety of spatial scales, including the county level. Like Domina, they find concentration to be generally increasing at the county level from the late 1990s. Interestingly, they find that births, deaths, and immigration all caused the index of concentration to increase, but net internal migration on its own would have led to deconcentration during the period.

1.4 Measurement: Disaggregating the Lorenz Curve

When the Lorenz curve is split into two pieces by the vertical line representing the Hoover index and the maximum difference between cumulative populations and cumulative areas, each of the two pieces may then be scaled and transformed into a Lorenz curve. Each of these has an associated Hoover index. If desired, each of these in turn may also be split into two pieces and Hoover indexes calculated for each of the new pieces; this process can continue to be repeated. Thus it is possible to examine inequality along portions of the Lorenz curve—in the application that is the focus here, we may examine population concentration for sets of regions that have a particular range of population density.

The Gini coefficient is always at least as high as the Hoover index. The excess is equal to twice the area between the Lorenz curve and the 45° line that lies beneath the triangle created by the end points of the Lorenz curve and the lower point of the vertical line associated with the Hoover index. The Gini coefficient for the full Lorenz curve may be calculated as the sum of the overall Hoover index and a weighted sum of the Hoover indexes associated with the disaggregation described above, where the weights are equal to the proportion of area (or population, in the case of application to income inequality) associated with that segment of the curve (Rogerson 2013).

We now begin with an illustration for the US population in 2010, at the county level. The Hoover index was 66.16 (after rounding to two decimal places); 16.31% of the population lived on 82.48% of the land (and, of course, the other 83.69% of the population lived on just 17.52% of the land). The former group consisted of 2166 low-density counties, and the latter group had 977 high-density counties. The low-density counties had an average of 17.29 people per square mile; the high-density counties averaged 417.5 people per square mile.

Each of these two groups may be examined as a separate subsystem. The line 0B in triangle 0AB in Fig. 1.1 represents what would be expected if all low-density counties had the same density. Similarly, the line BD in triangle BCD represents what would be expected if all high-density counties had the same population density. There is of course concentration *within* low-density and high-density subsystems, and this is captured by these two triangles.

Among the 2166 low-density counties, the 795 with the lowest density have 15.96% of the population on 65.73% of the land area ($H_L = 65.73 - 15.96 = 49.77$). Among the 977 high-density counties, 687 have 34.29% of the population on 76.95% of the area, for a Hoover index of $H_U = 76.95 - 34.29 = 42.66$.

These three breakpoints—one for the entire Unites States and the other two for low- and high-density subsystems—serve to divide the United States into four categories of population density (labeled 1 through 4 on Fig. 1.1). The coordinates of points A, B, C, and D, representing cumulative population and cumulative area, are given in Table 1.1. The table also reports the absolute values of the differences between cumulative area and cumulative population. Note that the highest of these absolute differences (when multiplied by 100) represents the greatest vertical distance between the Lorenz curve and the 45° line and is equal to the Hoover index of 66.16.

The triangles for the low-density and high-density subsystems are redrawn in Figs. 1.2 and 1.3, respectively. Part (a) of each figure is the triangle as it appears within the original Lorenz curve in Fig. 1.1; part (b) of each figure has scaled the triangle in part (a) to represent a Lorenz curve for the two-region subsystem that the figure represents.

The length of the maximum vertical distance in part (a) of Figs. 1.2 and 1.3 can be thought of scaled versions of the full Hoover indexes found in part (b). For Fig. 1.2a, the *y*-coordinate of point P is found as (0.1631/0.8248) (0.5422) = 0.1072. The vertical distance and scaled Hoover index is then 0.1072 − 0.0260 = 0.0812. In Fig. 1.3a, point P has a *y*-coordinate of 0.1631 + (0.9596 − 0.8248)(1 − 0.1631)/ (1 − 0.8248) = 0.8070. The vertical distance and scaled Hoover index is 0.8070 − 0.4501 = 0.3569.
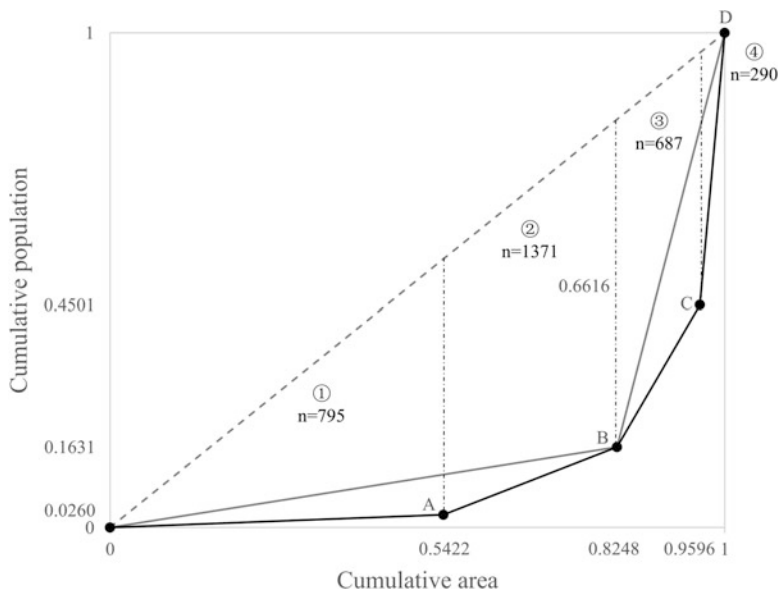
Fig. 1.1 Lorenz curve for US population: 2010

Table 1.1 Areas and populations for US counties, 2010

Population density category	Proportion of total population (p)	Proportion of total area (a)	Cumulative population (cp)	Cumulative area (ca)	Absolute difference ($cp - ca$)
1	0.0260	0.5422	0.026	0.5422	0.5161
2	0.1371	0.2824	0.1631	0.8248	0.6616
3	0.2970	0.1348	0.4501	0.9596	0.5095
4	0.5499	0.0404	1.0	1.0	0

The Gini coefficient (for the Lorenz curve in Fig. 1.1) may be calculated as a weighted sum of these scaled Hoover indexes, where the weights are the proportion of total land area that is associated with each section of the Lorenz curve. Thus

$$G = .6616 + (.8248)(.0812) + (.1752)(.3569) = .7911.$$

The second and third terms in the sum also represent the areas of triangles OAB and BCD in Fig. 1.1 (repeated as Figs. 1.2a and 1.3a), respectively.

More generally, suppose p_i is the proportion of total population found in subset i, and a_i is the proportion of total area found in subset i (where i is indexed from 1 to 4, corresponding to the sections in Fig. 1.1). Let cp_i and ca_i be the corresponding cumulative proportions. Furthermore, subsets 1 and 2 together comprise the low-density population and subsets 3 and 4 together comprise the high-density population (as determined by the system-wide Hoover index). The Hoover index for the entire system is equal to the absolute value of the difference between cp_2 and ca_2.
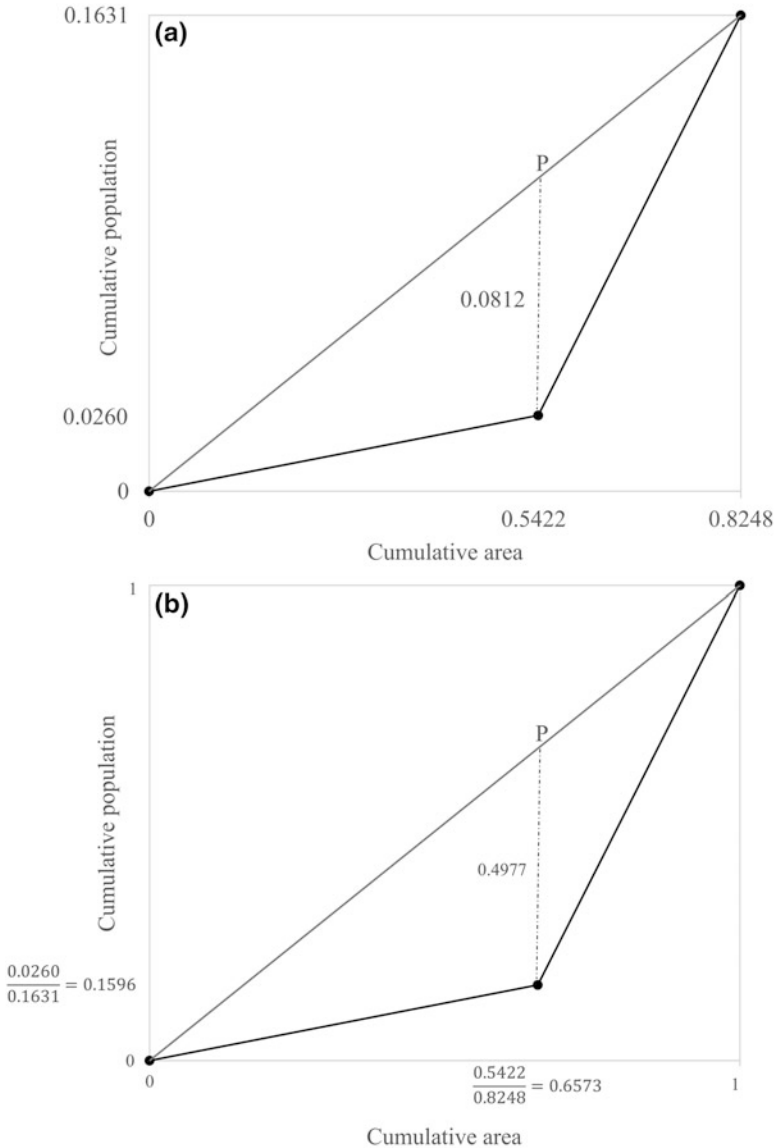
Fig. 1.2 Lorenz curve for low density US counties: 2010

The Hoover index for low-density subsystem in Fig. 1.2b (say H_L) is the absolute value of the difference between cp_1/cp_2 and ca_1/ca_2. The Hoover index for high-density subsystem in Fig. 1.3b, comprised of Sects. 3 and 4 in Fig. 1.1, is the absolute value of the difference between $(ca_3 - ca_2)/(1 - ca_2)$ and $(cp_3 - cp_2)/(1 - cp_2)$.
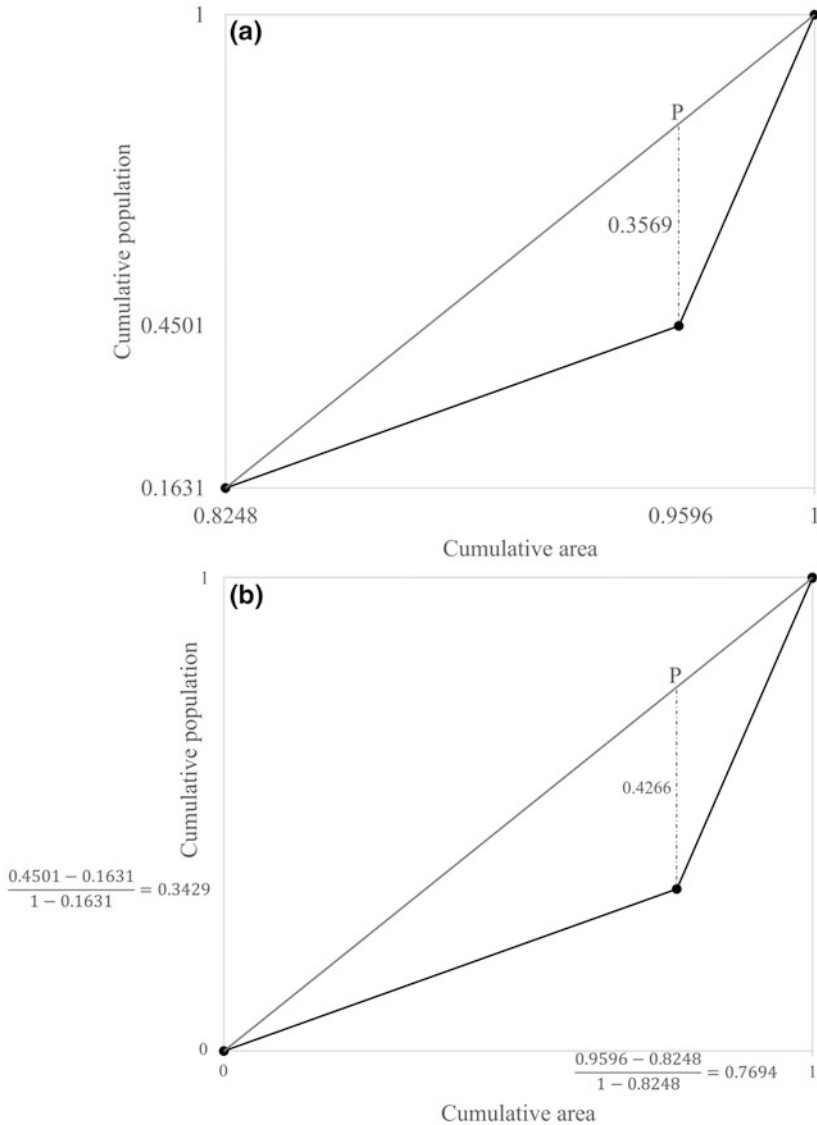
Fig. 1.3 Lorenz curve for high density US counties: 2010

The scaled Hoover indexes (H′), i.e., the heights of the two sub-triangles in Fig. 1.1, are found as the difference between the y-coordinate of point P and cp_1 in Fig. 1.2a and the difference between the y-coordinate of point P and cp_3 in Fig. 1.3a. The y-coordinate of point P in Fig. 1.2a is equal to $(cp_2/ca_2)(ca_1)$; in Fig. 1.3a it is equal to $cp_2 + (ca_3 - ca_2)(1 - cp_1)/(1 - ca_1)$. Thus

$$H'_L = (cp_2/ca_2)(ca_1) - cp_1$$
$$H'_U = cp_2 + (ca_3 - ca_2)(1 - cp_1)/(1 - ca_1) - cp_3$$

and this last equation can be simplified a bit by replacing the difference in cumulative proportions $(ca_3 - ca_2)$ by the proportion a_3 and $(cp_2 - cp_3)$ by $-p_3$:

$$H'_U = cp_2 + (a_3)(1 - cp_1)/(1 - ca_1) - cp_3$$
$$= (a_3)(1 - cp_1)/(1 - ca_1) - p_3$$

The Gini coefficient is then equal to $H + (ca_2) H_L' + (1 - ca_2) H_U'$.

1.5   Changes in Population Concentration: 2000–2015

In 2000, when the 2144 counties with lowest pop density are examined (out of a total of 3141 counties), we find that they contain 82.3% of the country's area, but just 16.7% of the population, leading to a Hoover index of 65.61. These counties have an average population density of 16.14 people per square mile; the remaining 997 high-density counties have an average population density of 374.81 people per square mile. During the first decade of this century then, there was an increase in population concentration at the county level, as the Hoover index rose from 65.61 to 66.16. By 2015, the US population was even more concentrated at the county level, with $H = 66.73$.

Among the 2144 low-density counties, the Hoover index was 49.31 in 2000, with the most rural of places (758 of the 2144 counties) containing 65.53% of the land area in this subset but just 16.22% of the population and having an average density of 3.99 people per square mile. The remaining $2144 - 758 = 1386$ counties had an average population density of 39.22 people per square mile. As a group, the low-density counties experienced a small increase in concentration, with the Hoover index rising from 49.31 to 49.77 during the decade. Concentration among the low-density counties continued during the first half of the next decade, with the Hoover index rising further to 49.87 by 2015.

Among the 997 high-density counties, the 719 least dense of these counties had 78.06% of the area but just 33.98% of the population (and an average population density of 163.04 people per square mile), leading to a Hoover index of 44.08 in 2000. The remaining high-density counties had an average population density of 1131.2 people per square mile. Within the high-density counties, deconcentration was experienced during the decade, with the Hoover index falling from 44.08 to 42.66. From 2010 to 2015, the index then rose slightly, to $H = 42.7$. These results are summarized in Table 1.2.

Table 1.2 Summary of changes in population concentration: 2000–2015

	2000	2010	2015
n counties	3141	3143	3143
Total pop	281,421,906	308,745,538	321,396,328
Total area	3,537,438	3,531,905	3,531,905
Pop density	79.56	87.42	91
H	65.61	66.16	66.73

Low-density counties

	2000	2010	2015
(*n*)	2144	2166	2185
Area	2,911,440 (0.823)	2,912,978 (0.8248)	2,923,256 (0.8277)
Population	46,977,694 (0.167)	50,370,014 (0.1631)	51,530,769 (0.1603)
Density	16.14	17.29	17.63

High-density counties

	2000	2010	2015
(*n*)	997	977	958
Area	625,498	618,927	608,649
Population	234,444,212	258,375,524	269,865,559
Density	374.81	417.46	443.38
Low-density counties *H*	49.31	49.77	49.87

Lower density portion of low density counties

	2000	2010	2015
(*n*)	758	795	816
Area	1,907,779 (0.6553)	1,914,827 (0.6573)	1,927,283 (0.6593)
Population	7,618,610 (0.1622)	8,038,937 (0.1596)	8,276,116 (0.1606)
Density	3.99	4.2	4.29

Higher density portion of low-density counties

	2000	2010	2015
(*n*)	1386	1371	1369
Area	1,003,661	998,151	995,973
Population	39,359,084	42,331,077	43,254,653
Density	39.22	42.41	43.43
High-density counties *H*	44.08	42.66	42.72

Lower density portion of high-density counties

	2000	2010	2015
(*n*)	719	687	672
Area	488,672 (0.7806)	476,267 (0.7695)	457,974 (0.7524)
Population	79,673,800 (0.3398)	88,603,570 (0.3429)	87,775,140 (0.3253)
Density	163.04	186.04	191.66

Higher density portion of High-Density Counties

	2000	2010	2015
(*n*)	278	290	286
Area	136,826	142,660	150,675
Population	154,770,412	169,771,954	182,090,419
Density	1131.15	1190.05	1208.5

Note: Areas are in square miles; densities are people per square mile

Table 1.3 Hoover index along four segments of Lorenz curve: 2000 and 2010

Year	2000	2010	2015
Low density portion of low-density counties			
H	43.42	43.43	43.68
% population	20.58	20.03	21.62
% area	64	63.45	65.3
n	286/758	310/795	330/816
High density portion of low-density counties			
H	17.93	18.65	18.95
% population	39.25	38.14	37.84
% area	57.18	56.78	56.79
n	730/1386	720/1371	722/1369
Low density portion of high-density counties			
H	20.25	20.95	20.7
% population	39.3	39.44	40.45
% area	59.54	60.39	61.15
n	425/719	412/687	399/672
High density portion of high-density counties			
H	30.98	29.57	29.96
% population	36.48	36.94	37.29
% area	67.45	66.51	67.24
n	160/278	169/290	161/286

Table 1.3 shows the results of further disaggregation, where there are four separate Hoover indexes associated with eight sections of the Lorenz curve. There has been little change in concentration over the period for the subset of lowest density counties. For the next level up the density hierarchy, there has been a steady increase in concentration between 2000 and 2015, with H increasing from 17.93 to 18.95 (and it is interesting to note that this subsystem of counties has the lowest H values, indicating relative uniformity in density). At the next step up the curve, there was also an increase in concentration between 2000 and 2010 and then a slight decrease during the first half of this decade. Finally, for the subset of counties with the highest densities, there was deconcentration during 2000–2010, followed by a slight increase in concentration during 2010–2015.

1.6 Attributing Change to Particular Counties

The Hoover index does not change when people move from a region on one side of the vertical line that defines the Hoover index to another region on that same side of the line. Migration will only cause change in the index when people move from a

Table 1.4 Counties contributing to increasing concentration within the high density portion of the low density counties

State	County	Population 2000	Population 2015	Percent change, 2000–2015
Arizona	Pinal County	179,277	406,584	126.8
Colorado	Mesa County	116,255	148,513	27.7
Colorado	Summit County	23,548	30,257	28.5
Florida	Baker County	22,259	27,420	23.2
Florida	Wakulla County	22,863	31,535	37.9
Georgia	Bacon County	10,103	11,299	11.8
Georgia	Candler County	9577	10,886	13.7
Georgia	Lanier County	7241	10,312	42.4
Georgia	Long County	10,304	17,731	72.1
Georgia	Pulaski County	9588	11,396	18.9
Hawaii	Hawaii County	148,677	196,428	32.1
Minnesota	Mille Lacs County	22,330	25,788	15.5
Montana	Missoula County	95,802	114,181	19.2
Oklahoma	Marshall County	13,184	16,232	23.1
Oregon	Deschutes County	115,367	175,268	51.9
Oregon	Hood River County	20,411	23,137	13.3
Texas	Austin County	23,590	29,563	25.3
Texas	Burnet County	34,147	45,463	33.1
Texas	Kendall County	23,743	40,384	70.0
Texas	Polk County	41,113	46,972	14.3
Texas	San Jacinto County	22,246	27,413	23.2
Texas	Somervell County	6809	8739	28.3
Utah	Washington County	30,373	34,765	14.4

region on one side of the vertical line to a region on the other side of the line. When two Lorenz curves are compared, there may be regions that are on one side of the vertical line at one point in time and on the other side of the line at the next point in time. A region on the left side of the vertical line in 1 year will contribute to increased concentration if it is on the right side of the line in a subsequent year.

To illustrate, here we examine the set of counties that had a density of about 40 persons per square mile in 2000 and find those that contributed to the increased concentration that was observed over the next 15 years. In 2000, there were 1386 counties in Sect. 2 of the Lorenz curve (when it is divided into four sections, in order of increasing density); there were 1369 counties in that section in 2015. In 2000, these counties had a population density of 39.22 persons per square mile, and they ranged in population density from 16.2 to 79.4 people per square mile. By 2015, this section of the curve contained counties with an average population density of 43.43 persons per square mile and densities ranging from 17.7 to 90.8 persons per square mile. The 23 counties listed in Table 1.4 were among the low-density counties in this

portion of the Lorenz curve in 2000 but were in the high density portion in 2015. They therefore contributed to the increased concentration in this subsystem. Many of these grew rapidly and are easily recognizable as amenity-laden places. Others, such as Bacon County, Georgia, did not grow particularly fast but were close to the vertical line in the Lorenz curve of 2000 (near the mean density of 39.22 persons per square mile); even a small amount of growth was sufficient to put them on the right-hand side of the vertical line in the 2015 Lorenz curve.

1.7   Discussion and Summary

The Hoover index of concentration has a long history of application in population geography and demography. It has received particularly widespread use in the study of rural population change since the 1970s.

Like the Hoover index, the Gini coefficient may be interpreted in relation to the Lorenz curve. It may also be derived as the sum of the Hoover index, plus weighted, scaled Hoover indexes associated with partitions of the Lorenz curve constructed by dividing the curve at the vertical line associated with the Hoover index. This approach was used here to show how population concentration has changed during the period 2000–2015. Overall increases in concentration were accompanied by concentration for many population density levels, but deconcentration occurred for that set of counties comprising the highest density levels.

It would be interesting to assess the temporal change in Hoover indexes for these subsystems for various demographic components of change. Rogerson and Plane (2013) found increasing concentration for births, deaths, and international migration and *deconcentration* for internal migration. Whether these trends hold at all levels of population density is unknown. Rural areas, e.g., tend to have a relatively older age structure, and the relatively large number of deaths and small number of births act to increase population concentration in the more dense places among them. This effect might, for example, be less apparent or nonexistent at higher population densities. It would be particularly interesting to know, e.g., whether net internal migration is leading to deconcentration at all levels of population density or whether only a certain portion of the density hierarchy is witness to that trend.

Finally, it could also be enlightening to examine more closely the geographic distribution of these changes. The focus here has been on groupings of counties by population density. These groupings could easily be mapped, as could the counties that increased or decreased in population density sufficiently to contribute to a change in the Hoover index. Similarly, the analysis described above could be carried out for regional divisions of the United States.

References

Berry BJL (1980) Urbanization and counterurbanization in the United States. Ann Am Acad Pol Soc Sci 451(1):13–20

Bourne LS (1980) Alternative perspectives on urban decline and population deconcentration. Urban Geogr 1:39–52

Ceriani L, Verme P (2012) The origins of the Gini index: extracts from Variabilitia e Mutabilita (1912) by Corrado Gini. J Econ Inequal 10:421–443

Cochrane SG, Vining DR Jr (1988) Recent trends in migration between core and peripheral regions in developed and advanced developed countries. Int Reg Sci Rev 11:215–243

Dalton H (1920) The measurement of the inequality of incomes. Econ J 30:348–461

Domina T (2006) What clean break? Education and nonmetropolitan migration patterns, 1989–2004. Rural Sociol 71(3):373–398

Duncan OD, Cuzzort R, Duncan B (1961) Statistical geography. Free Press, Glencoe

Florence PS (1953) The logic of British and American industry: a realistic analysis of economic structure and government. Routledge and Kegan Paul, London

Florence PS, Wensley AJ (1939) The location of industry in Great Britain. London

Frey WH (March 26, 2018) US population disperses to suburbs, exurbs, rural areas, and "middle of the country" metros. https://www.brookings.edu/blog/the-avenue/2018/03/26/us-population-disperses-to-suburbs-exurbs-rural-areas-and-middle-of-the-country-metros/. Last accessed 2 Dec 2018

Fuguitt G (1985) The nonmetropolitan population turnaround. Annu Rev Sociol 11:259–280

Geyer HS, Kontuly T (1993) A theoretical foundation for the concept of differential urbanization. Int Reg Sci Rev 15(2):157–177

Gini C (1912) Variabilit'a e Mutabilit'a. Contributo allo studio delle distribuzioni e delle relazioni statistiche, Studi economico-giuridici Anno III, Parte II. Facolt'a di giurisprudenza della Regia Universit'a di Cagliari. Cuppini, Bologna

Gini C (1914) Sulla Misura della Concentrazione e della Variabilit'a dei Caratteri. Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti LXXIII(II):1203–1248

Gini C (1921) Measurement of inequality of incomes. Econ J 31:124–126

Gini C (2005) On the measurement of concentration and variability of characters. METRON - Int J Stat LXIII(1):3–38

Giorgi GM (2014) A couple of good reasons to translate papers of the Italian statistical tradition. Metro 72(1):1–3

Gordon P (1979) Deconcentration without a 'clean break'. Environ Plan A 11(3):281–289

Hall SA, Kaufman JS, Ricketts TC (2006) Defining urban and rural areas in U.S. epidemiological studies. J Urban Health 83(2):162–175

Hoover EM (1936) The measurement of industrial localization. Rev Econ Stat 18:162–171

Hoover E (1941) Interstate redistribution of population: 1850–1940. J Econ Hist 1:199–205

Johnson KM, Cromartie JB (2006) The rural rebound and its aftermath. In: Kandel WA, Brown DL (eds) Population change and rural society. Springer, Dordrecht, pp 25–49

Lichter DT (1985) Racial concentration and segregation across US counties, 1950–1980. Demography 22(4):603–609

Lichter DT, Fuguitt GV (1982) The transition to nonmetropolitan population deconcentration. Demography 19(2):211–221

Lichter DT, Johnson KM (2006) Emerging rural settlement patterns and the geographic redistribution of America's new immigrants. Rural Sociol 71(1):109–131

Long J (1981) Population deconcentration in the United States. Department of Commerce, US Government Printing Office, Washington, DC

Long L, Nucci A (1997a) The 'clean break' revisited: is US population again deconcentrating? Environ Plan A 29(8):1355–1366

Long L, Nucci A (1997b) The Hoover index of population concentration: a correction and update. Prof Geogr 49(4):431–440

Lorenz MO (1905) Methods of measuring the concentration of wealth. Q Publ Am Stat Assoc ix:209–20I

McCarthy KF, Morrison PA (1977) The changing demographic and economic structure of nonmetropolitan areas in the United States. Int Reg Sci Rev 2(2):123–142

Morrill R (1979) Stages in patterns of population concentration and dispersion. Prof Geogr 31 (1):55–65

Morrill R (1980) The spread of change in metropolitan and nonmetropolitan growth in the United States, 1940–1976. Urban Geogr 1(2):118–129

Parker K, Horowitz JM, Brown A, Fry R, Chon D, Igielnik R (May 22, 2018) What unites and divides urban, suburban, and rural communities. Pew Research Report. http://www.pewsocialtrends.org/2018/05/22/demographic-and-economic-trends-in-urban-suburban-and-rural-communities/ Last accessed 2 Dec 2018

Pietra G (1915) Delle relazioni tra gli indici di variabilità (Nota I)" *Atti del Reale Istituto Veneto di Scienze, Lettere e Arti. 1915, Vol. LXXIV, Part I, pages 775–792*

Pietra G (2014) On the relationships between variability indices (note I) (translated from the 1915 article) in *Metron* 72(1):5–16

Plane DA, Mulligan GF (1997) Measuring spatial focusing in a migration system. Demography 34 (2):251–262

Plane DA, Henrie CJ, Perry MJ (2005) Migration up and down the urban hierarchy and across the life course. Proc Natl Acad Sci 103(43):15313–15318

Rogers A, Raymer J (1998) The spatial focus of US interstate migration flows. Int J Popul Geogr 4:63–80

Rogers A, Sweeney S (1998) Measuring the spatial focus of migration patterns. Prof Geogr 50 (2):232–242

Rogerson P (2013) The Gini coefficient of inequality: a new interpretation. Lett Spat Resour Sci 6 (3):109–120. https://doi.org/10.1007/s12076-013-0091-x

Rogerson PA, Plane DA (2013) The Hoover index of population concentration and the demographic components of change: an article in memory of Andrew M. Isserman. Int Reg Sci Rev 36(1):97–114

Schutz R (1951) On the measurement of income inequality. Am Econ Rev 41:107–122

US Bureau of the Census (2012) United States summary: 2010. 2010 census of population and housing, population and housing unit counts, CPH-2-5. U.S. Government Printing Office, Washington, DC

Vinci F (1934) Manuale di Statistica, I (Bologna), pp 113–118

Vining D, Strauss A (1977) A demonstration that the current deconcentration of population in the United States is a clean break with the past. Environ Plan A 9:751–758