Urban Studies, Vol. 44, No. 10, 1889–1904, September 2007

Urban Land Area and Population Growth: A New Scaling Relationship for Metropolitan Expansion

Julian D. Marshall

[Paper first received, December 2005; in final form, November 2006]

Summary. In most metropolitan regions throughout the globe, urbanised land area is increasing to accommodate increasing population size. This article provides a simple yet powerful mathematical description of this urban expansion. Specifically, the following scaling relationship is proposed: land area (A) increases proportionally to population size (P) raised to a power (n)—i.e. $A \propto P^n$. During 1950–2000, this relationship is found to hold well for US Census urban areas (UAs) with a greater than 10 per cent increase in population. Values for the parameter n vary among UAs, with a central tendency value of $\sim$ 2, suggesting that, on average, newcomers to urban areas occupy about twice the land area per capita of existing residents. If n were exactly equal to 2, then the parameter group $P/\sqrt{A}$ (called 'linear population density', or LPD) would be constant over time. LPD (units: people per metre) is the number of people in a metre-wide strip across an urban area. LPD is distinct from, and behaves somewhat differently than, population density. Distributions of LPD values among US UAs during 1950–2000 show surprisingly little variability over multidecade time-scales. For example, from 1950 to 2000, average population, land area and population density changed by more than a factor of 2, but average LPD changed less than 10 per cent. Few, if any, other attributes of urban form have remained so constant during this half-century time-period. International data corroborate the finding that LPD distributions are roughly constant over multidecadal time-scales. These results suggest an underlying pattern to how people arrange themselves within and among urban areas. For US UAs, rank–size rules similar to the generalised version of Zipf's rule hold for population, land area, LPD and population density. LPD is an important predictor of the emissions-to-inhalation relationship for motor vehicle emissions. Results presented here are important for theoretical, practical and empirical investigations of urban form and of how urban areas expand over time.

Introduction

Urban populations are increasing rapidly. In approximately 2007, for the first time in history, more people will live in urban than rural areas (United Nations, 2000b). In coming decades, urban populations are expected to double to ~6 billion in the next 50 years, while rural populations remain constant (at ~3 billion) or decline. Population growth affects urban areas in many ways, from infrastructure requirements and their environmental impacts, to new patterns of social interactions and

Julian D. Marshall is in the Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive S.E., Minneapolis, MN 55455–0116, USA. Fax: 612 626 7750. E-mail: julian@umn.edu. *The author thanks Mike Toffel for fruitful discussions on the topics contained here, also three anonymous reviewers and the journal editor Andrejs Skaburskis for helpful comments and suggestions. Funding from the Bridge Program at University of British Columbia, and start-up funds from University of Minnesota, are gratefully acknowledged.*

0042-0980 Print/1360-063X Online/07/101889–16 © 2007 The Editors of Urban Studies
DOI: 10.1080/00420980701471943

changes in the regional economy (Bettencourt *et al.*, 2007; Rosser, 1980). The issue of where urban growth occurs is important to understanding this growth and to predicting its broad influences. In the US, urban areas are growing (albeit not as rapidly as the global rates) and urban land area is expanding faster than urban population size, leading to a decline in average urban population density.

This article explores how urban land area expands over time in response to growth in urban population, a topic of theoretical and practical importance. The paper is comprised of three sections. First, I propose a scaling rule relating changes in urban land area to changes in urban population size. I test the rule's validity using comprehensive panel data for US urban areas during 1950–2000. Next, I introduce a new urban form metric called 'linear population density' (LPD), show that the distribution of LPD values for US urban areas is relatively constant over multiple decades and present evidence for population density and LPD rank–size rules. Finally, I discuss implications of these findings for a specific environmental health issue: human inhalation of motor vehicle emissions in urban areas. This article documents previously unobserved spatial patterns in urban growth. Topics such as elucidating causal mechanisms underlying these patterns and discussing whether planners can or should seek to modify these extant trends are left to future research.

Scaling Rule

Background

Scaling relationships describe in general terms how two or more attributes are related. As a straightforward example, the volume of an object scales with (i.e. is proportional to) the cube of linear size: doubling the radius of a sphere would increase its volume by a factor of 8. Thus far, two scaling rules address urban form. One rule focuses on the distribution of population sizes among urban areas. Zipf (1949) hypothesised that the population of any specific urban area (P_i) scales with the population rank (i) of that area, according to $P_i = P_1/i$, where P_1 is the population of the largest urban area. For example, the population of the second-largest urban area equals half the population of the largest urban area; the population of the third-largest area equals one-third the population of the largest area, etc. Zipf's rule can be generalised as $P_i = Ki^{\alpha}$, where K is a constant close in value to P_1, and α has a value close to -1 (Pumain, 2003). Most research supports the generalised version of this 'rank–size rule' (Gabaix, 1999; Black and Henderson, 2003; Ioannides and Overman, 2003; Urzua, 2000).

A second scaling rule focuses on how population density varies within an urban area. Clark (1951) proposed the monocentric exponential decay equation $D(r) = D_0 e^{-\lambda r}$, where $D(r)$ is the population density (km^{-2}) at a specific distance, r (km), away from the city centre and D_0 (km^{-2}) and λ (km^{-1}) are empirical coefficients. In theory, D_0 is the population density at the city centre (i.e. at $r = 0$). In practice, city centres are often park or commercial districts, where residential population density is less than empirically determined values for D_0. Subsequent research has applied this equation in hundreds of urban areas (Edmonston *et al.*, 1985), often estimating λ but not D_0 (Mills and Tan, 1980); proposed alternative, often polycentric, density gradient equations, including polynomials (Bunting *et al.*, 2002), inverse power relationships (Batty and Kim, 1992; Stern, 1993), exponential relationships (Anas *et al.*, 2000; Murakami *et al.*, 2005; Newling, 1966) and a cubic-spline (Anderson, 1985; Skaburskis, 1989; Zheng, 1991); and, explored causes and consequences of density gradients (Anas *et al.*, 1998).

Building on the two rules above, this article proposes a new scaling rule, regarding how urban areas expand over time. Findings presented below provide a mathematical description of where populations locate within and among urban areas over time. The current lack of a scaling rule for expansion of urban areas is a significant gap in the literature, especially given the expected growth in urban populations.

Scaling over Time between Land Area and Population Size

The scaling relationship proposed here relates, for a given urban area, land area occupied (A; units: km^2) and population size (P). It is posited that area expands to accommodate increases in population size according to equation (1)

$$A \propto P^n \tag{1}$$

where n is an empirical constant and '$\propto$' means 'is proportional to'. Here, the proportionality (scaling) is over time for a specific urban area and not among urban areas. Equation (1) can equivalently be written as

$$\frac{A_2}{A_1} = \left(\frac{P_2}{P_1}\right)^n \tag{2}$$

where, for a given urban area, A_1 and A_2 are the land areas at times 1 and 2 respectively, and P_1 and P_2 are the population sizes at times 1 and 2 respectively. This one-parameter equation captures the essential relationship being described, but an inherent limitation is if land area varies while population is constant (or nearly constant), then the estimated magnitude for n is infinite (or very large). Equations (1) and (2) cannot be proved or disproved: an estimate for n can be derived for any urban area with known initial and final population and area, so long as land area, population and population growth are non-zero. Instead, we consider whether this equation offers a *useful* description of the phenomena being investigated.

If n were constant, the derivative of equation (1) indicates that

$$\frac{dA}{A} = n\frac{dP}{P} \tag{3}$$

where dA and dP represent small changes in area and population respectively. The parameter n is the population–area elasticity. For example, in a city with $n = 3$, a 2 per cent increase in population yields a $\sim$6 per cent increase in area. In this case, on average, newcomers to the urban area occupy about three times the per capita land area as do current residents.

Empirical Estimates of the Parameter 'n'

Estimating values for n requires time-series data (i.e. values for one urban area over time) or panel data (i.e. values for multiple urban areas over time) rather than cross-sectional data (i.e. values for multiple urban areas at a single time). The data must provide urban populations and land areas based on urban boundaries (for example, determined by population density and continuous areas of growth) rather than political boundaries. Appropriate panel data are available for US urban areas: the US Census delineates 'urban areas' based on the population density of a census block or block group being greater than 1000 $mile^{-2}$ (386 km^{-2}) and the population density of surrounding census blocks being greater than 500 $mile^{-2}$ (193 km^{-2}) (US Census, 2004). Urban areas (UAs) must have a population size of 50 000 or greater. UAs are delineated at the start of each decade. Further information on UAs is available elsewhere (US Census, 2004). Uncertainty estimates for the UA data are unavailable. US Census undercount rates declined from roughly 4 per cent in 1950 to less than 1 per cent in 2000 (CNSTAT, 2004; Mulry, 2006; Robinson *et al.*, 1993). Undercount rates are typically lower in urban than rural areas (Hogan and Robinson, 2000).

The year 1950 is chosen as the starting-point for this analysis because the US Census definition of an urban area changed significantly between the 1940 and 1950 censuses, but subsequent modifications have been minor (US Census, 1992). The number of UAs increases over time because of population growth. Typically, once an area becomes an 'urban area', it stays such. (The two exceptions are Concord, NC, and Danville, IL, which were US census UAs in 1980 and 2000, but not 1990. These two areas were excluded from my analysis.)

Urban areas are grouped in Table 1 based on the decade when the Census first classified the area as an urban area. Group 1 contains

Table 1. Year 2000 values for urban form attributes for six urban area (UA) groups

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	All six groups combined
Year area became a US census urban area (UA)	1950 or before	1960	1970	1980	1990	2000	–
Number of censuses with data during 1950–2000	6	5	4	3	2	1	–
Number of UAs in group	142	50	34	63	72	92	453
Portion of US population in group (percentage)	56	4	2	2	3	3	70
Population per UA							
Mean[a]	1 080 000	213 000	165 000	104 000	114 000	92 100	425 000
Median	394 000	120 000	143 000	85 200	84 800	59 700	118 000
Land area per UA (km^2)							
Mean[a]	1470	350	320	240	260	200	640
Median	810	250	230	180	190	150	250
Population density							
Mean (km^{-2})[a]	610	600	580	500	490	500	550
Coefficient of variability (percentage)[b]	40	42	38	44	39	43	42
Median (km^{-2})	540	540	540	410	430	430	490
Population-weighted mean (km^{-2})	870	730	600	520	490	540	810
Geometric mean (km^{-2})	570	550	540	470	460	460	510
Geometric standard deviation	1.4	1.5	1.4	1.5	1.4	1.5	1.5
Linear population density							
Mean (m^{-1})[a]	22	11	9.3	6.9	7.0	6.4	12
Coefficient of variability (percentage)[b]	97	63	37	37	35	48	120
Median (m^{-1})	14	8.4	8.8	6.4	6.4	5.4	7.8
Population-weighted mean (m^{-1})	61	17	11	8.0	8.6	8.9	51
Geometric mean (m^{-1})	17	9.2	8.8	6.6	6.7	5.9	9.2
Geometric standard deviation	2.0	1.7	1.4	1.4	1.4	1.4	1.9

[a]'Mean' refers to arithmetic mean.
[b]Coefficient of variability is the arithmetic standard deviation divided by the arithmetic mean.

areas that were classified as UAs in the 1950 census; Group 2 contains areas first classified as UAs in 1960, and so on through Group 6, which contains areas first classified as UAs in the most recent census in 2000. Each group represents a unique set of UAs.

All six groups differ from each other, but Group 1 stands apart from the other groups for the following reasons. A majority of US residents (56 per cent) live in Group 1, compared with only 2–4 per cent of the US population in each of the remaining groups. Group 1 contains 36 of the 37 UAs where the year 2000 population exceeded 1 million people. (The exception is Las Vegas, Nevada, which was first classified as an UA in 1960.) Mean and median values for urban population size and land area are ~3–10 times larger for Group 1 than for the remaining groups.

The maximum duration of data for each UA is used in equation (2) to estimate the parameter *n* (for example, for Group 1, changes between 1950 and 2000 are considered; for Group 2, changes between 1960 and 2000 are considered, etc.). Equation (1) is only proposed to apply for modest or rapid population growth. Thus, among the 361 UAs (see Table 1) with data for two or more censuses, I removed the 51 UAs (14 per cent) for which population either increased less than 10 per cent or declined (again employing the maximum duration of data for each UA). For the remaining 310 UAs, Figure 1 presents separately the distribution of *n* values for each of the UA groups (i.e. depending on the census year the territory became an UA). Among the five distributions in Figure 1, central tendencies (means and medians) are in the range 1.8–3.0 and interquartile values (i.e. spanning the 25th and 75th percentiles) are in the range 1.3–3.7. The overall mean and median values for *n* are 2.7 and 2.3 respectively. For the roughly two-thirds of UAs in this figure that experienced more than a 50 per cent increase in population size, the mean and median values for *n* are 2.1 and 2.0 respectively; for the remaining UAs in Figure 1 (i.e. the roughly one-third of UAs with a 10–50 per cent increase in population size), mean and median *n* values are 3.7 and 3.3 respectively. (Median and mean population size increases are 70 per cent and 130 per cent respectively.) Consistency among the distributions in

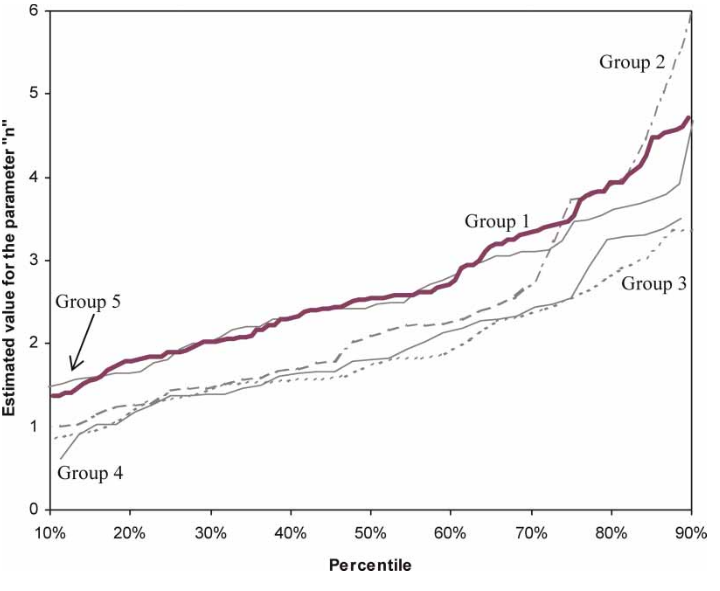

Figure 1. Distribution of values estimated for the parameter *n* in equation (1), for each of the five urban area groups with two or more censuses. The line for Group 1 in bold because the total population for this group of UAs is significantly larger than for the other groups.

Figure 1 suggests that the approach employed of estimating n based on the maximum possible duration for each group (i.e. 50 years for Group 1, 40 years for Group 2, etc.) is appropriate.

In conclusion, equation (1) provides a useful description of the data, with distributions of n values given in Figure 1. The median value for n is ~ 2. Typical n values are larger for modest population growth than for significant population growth (for UAs with 10–50 per cent and >50 per cent population increase, median n values are ~ 3 and ~ 2 respectively). As indicated above, n values, which vary among UAs, provide information about the population–area elasticity and about the typical land area occupied by urban newcomers relative to existing urban residents. For $n = 2$ (on average, urban newcomers occupy twice the land area of current residents), about half of the UA's land area expansion over time is attributable to population growth and half is attributable to rising per capita land consumption.

As a comparison with these results, two alternative approaches were also used to estimate values for n. The first approach employs the same data, but rather than investigating each UA separately, instead considers the total (i.e. cumulative) population and land area for each UA group. As before, the maximum possible duration is investigated for each group. Resultant values for n range from 1.6 to 2.3 and average 1.9. The second approach employs the exponential urban population density gradient proposed by Clark (1951). This analysis requires specifying the threshold population density separating 'urban' from 'not urban'. The US census value of 1000 mile^{-2} is used here. Relevant equations are in the Appendix. Using parameter values given by Edmonston *et al.* (1985) for 1951–1976 yields estimated n values of 1.8 for US urban areas with more than 1 million people. Estimated values for n are between 1.9 and 2.1 for US urban areas with between 250 000 and 1 000 000 people, and for Canadian urban areas. These values are reasonably robust to small changes in input parameters. (In some cases, application of the Appendix equations to exponential population density gradient data yields results that are highly sensitive to input parameters—for example, a 10 per cent change in the input values can yield factor-of-2 or larger changes in the n estimate. In this situation, estimates for n would not be considered robust.) Thus, both alternative approaches suggest the value ~ 2 for n, which is consistent with central tendencies of the n distributions presented earlier.

Linear Population Density

The preceding analyses of US census data and exponential population gradient data from Edmonston *et al.* (1985) suggest a central tendency value of $n \approx 2$, especially for large changes (more than 50 per cent increase) in population size. In the case where n is exactly equal to 2, an implication of equation (1) is that the parameter group $PA^{-0.5}$ ('linear population density', or LPD) would remain constant over time. It is worthwhile to investigate whether LPD is, in fact, approximately constant over time. If that were the case, LPD might merit further consideration as an urban planning tool—for example, when making long-term urban growth forecasts.

The linear population density of an urban area is the urban population divided by a linear measure (length or width) of the urban area, which I take here to be the square root of land area. LPD has an intuitive meaning: a value of 50 people per metre, for example, would mean that a metre-wide strip of land across an urban area contains 50 people. Figure 2 illustrates population density and LPD.

LPD is not merely a restatement of population density. For example, when ranking US states by population density and by LPD, some states have a similar rank in both lists; others change rank significantly. New Jersey ranks first in both population density and LPD; Alaska ranks last in both. In contrast, Rhode Island is ranked second for population density, but sixteenth for LPD; California is twelfth for population density, but second for LPD. US counties show similar patterns:

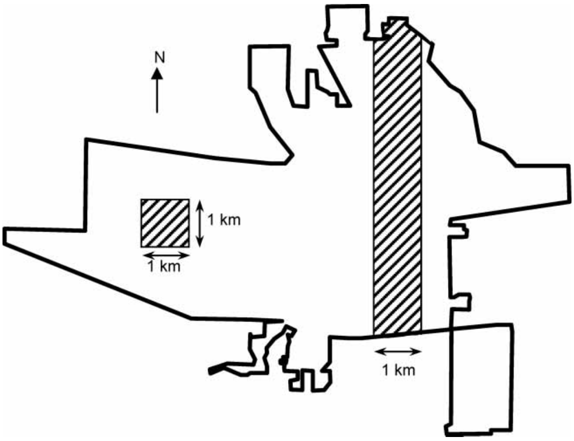

Figure 2. Outline map for an illustrative urban area: Ames, Iowa (population, P, is 50 700; land area, A, is 40.9 km^2). Average population density (PA^{-1}) is 1240 km^{-2} and average linear population density (LPD)—the number of people in a strip of land extending fully across the urban area ($PA^{-0.5}$)—is 7930 km^{-1}. While population density and LPD are similar, LPD is not merely a restatement of population density. Mathematically, LPD is distinct from, and behaves somewhat differently than, population density (see Figures 3–6).

population density and LPD ranks are similar for some counties and significantly different for others.

The main reason LPD differs from population density is straightforward: LPD depends less on area (A) than population density does (i.e. LPD is proportional to $A^{-0.5}$; density is proportional to A^{-1}). Thus, urban areas with a smaller population size tend to have a smaller LPD value, but not necessarily a smaller population density. As illustration, consider Ames, IA ($P = 50\ 700$; $A = 40.9\ km^2$), New Orleans, LA ($P = 1009000$; $A = 825\ km^2$), and New York, NY ($P = 1\ 780\ 000$; $A = 13\ 970\ km^2$), in year 2000. Population densities for these three UAs are comparable ($\sim$1250 km^{-2}) but LPD varies: LPD for New York (151 m^{-1}) is 4 times larger than for New Orleans (35 m^{-1}) and 19 times larger than for Ames (8 m^{-1}). In terms of population and land area, New York is $\sim$17 times larger than New Orleans and $\sim$346 times larger than Ames.

Although not incorporated in this analysis, another potential difference between population density and LPD relates to the shape and aspect ratio (i.e. length divided by width) of an urban area. Once an urban area is defined spatially, land area and population density are relatively unambiguous concepts. In contrast, LPD depends on which linear measure of an urban area is employed: mean width, maximum width, or other measures. This paper employs square root of land area. Other approaches may yield different results. Like population density, LPD for an urban area may vary spatially and temporally.

Current and Historical US Values for Linear Population Density

Year 2000 LPD values for US census UAs are presented in Table 1. Population-weighted year 2000 values for LPD and population density are 51 m^{-1} and 810 km^{-2} respectively. In general, variability among UA groups is greater for LPD than for population density.

Figure 3 presents normalised mean values of four urban form attributes for the six groups in Table 1. For all UA groups, changes in LPD over multiple decades are small (less than 10 per cent) and significantly less than changes in population, area and population density. The consistency over time of LPD is surprising. To the author's knowledge, no other attribute of urban form has remained as constant over the same five-decade period.

Figure 4 presents distributions of population density and LPD during 1950–2000 for the 142 UAs in Group 1. Over time, population density has declined but the LPD distribution retains curve shape and magnitude. Again, the degree of consistency over five decades is surprising. Geometric means (GM; units: m^{-1}) and geometric standard deviations (GSD; dimensionless) for the LPD distributions in Figure 4 are 17–19 and 1.8–2.0 respectively. Over time, GM has declined and GSD has increased, but changes are small. These results suggest the following prediction: LPD values during the next

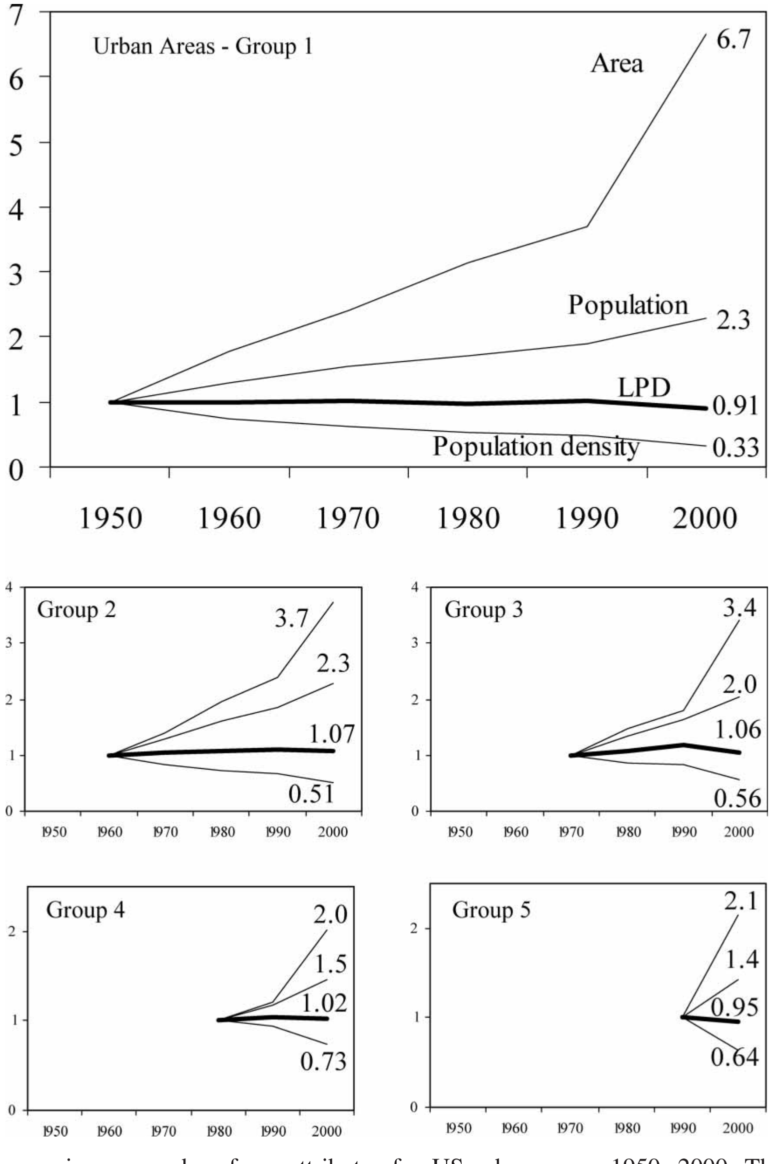

Figure 3. Changes in mean urban form attributes for US urban areas, 1950–2000. The large graph presents values for Group 1; the remaining four charts show values for Groups 2–5. Values are normalised to the first year of data (i.e. year 1950 for Group 1, year 1960 for Group 2, etc.) and the numerical label is the year 2000 value for each line. For all five plots, the four lines represent (in order, from top to bottom) area, population, linear population density (LPD) and population density. For all plots, population, area and population density change significantly, but LPD remains relatively constant.

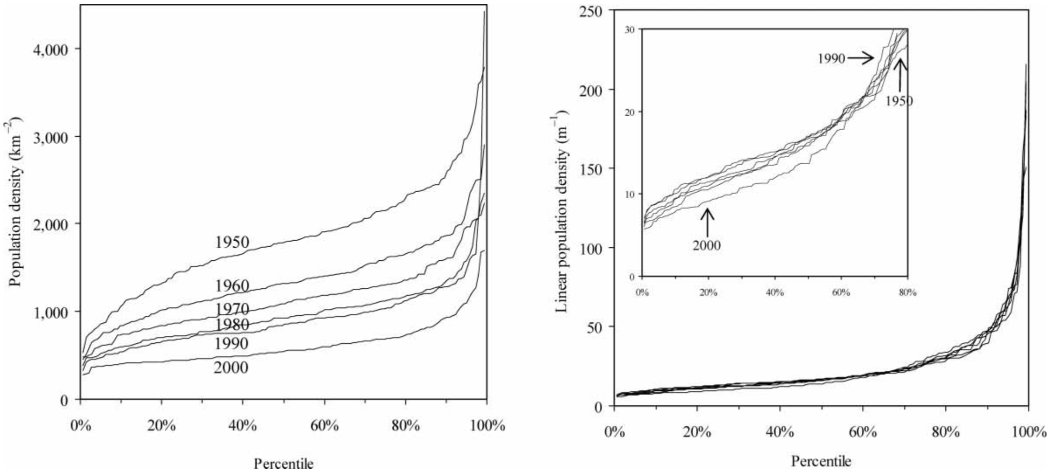

Figure 4. Distribution of population density and linear population density (LPD) for the six census years (1950–2000) for Group 1 urban areas. While population density distributions have declined over time, the LPD distribution has not changed significantly.

three censuses (years 2010, 2020 and 2030) are expected to have GM and GSD of 14–17 m^{-1} and 2.0–2.3 respectively.

As illustration of how LPD can remain constant over time while population density declines, consider the first UA alphabetically in Group 1: Akron, Ohio. During 1950–2000, Akron's population increased by 55 per cent (from 370 000 to 570 000 people) and its land area increased by a factor of 3.1 (from 255 to 797 km^2), yielding a factor of 2.0 reduction in population density (from 1440 to 715 km^{-2}). Square root of land area increased by 80 per cent (from 16.0 to 28.2 km) and LPD declined by only 12 per cent (from 23.0 to 20.2 m^{-1}). The value for n in equation (1), based on these data for Akron during 1950–2000, is 2.6. As Akron's population grew over time, each new resident on average occupied 2.6 times more land than existing residents. An average metre-wide strip of land across Akron, which increased in length from 16.0 km to 28.2 km during 1950–2000, contained roughly the same number of people in 2000 as in 1950 ($\sim$22 people), but those 22 people were living at a decreasing population density.

For individual UAs, LPD changes over time, either up or down. For Akron, OH, the change was minor—only 12 per cent during 1950–2000. In the same period, LPD (m^{-1}) increased by a factor of 3 for San Jose, CA (from 14 to 47) and decreased by a factor of 2 for Pittsburg, PA (from 60 to 29). Among the Group 1 UAs, the median change in LPD during 1950–2000 is 30 per cent, with 70 per cent of UAs experiencing a decline and 30 per cent of UAs experiencing an increase.

Thus, the overall LPD distribution is (nearly) unchanging, while individual UAs move up or down the distribution. Consistency over time in LPD distributions in Figure 4 results from population mobility both within and among UAs. Results presented here offer a novel and straightforward mathematical description of this long-term mobility.

Rank–Size Rules for Population, Area, Population Density and LPD

The four main urban form attributes investigated here—population, area, population density and LPD—conform well to rank–size rules that mimic the generalised form of Zipf's rank–size rule. For example, the generalised rank–size rule for population density (ρ) is $\rho_i = Ki^{\alpha}$, where ρ_i is the population density for the urban area with population density ranking i, and K and α are empirical constants. Generalised rank–size rules for the other three urban form attributes are analogous. Figure 5 presents population density and LPD values for Group 1 UAs; a straight line in these figures would indicate perfect conformity to a generalised rank-size rule. Best-fit lines for Figure 5, and for analogous figures for population and land area (not shown), yield the rank–size rule parameters for Group 1 UAs provided in Table 2. The R^2 values are close to 1, indicating that the data fit well to generalised rank–size rules. For the year 1950 population data, $\alpha \approx -1$ and K is approximately equal to the dataset largest value (15 million versus 12 million), indicating support for the original (non-generalised) rank–size rule. In contrast, LPD and population density data conform well to the generalised rank–size rule but not the original rule.

Values for n and LPD Internationally

The analyses above consider only US (and, in one analysis, Canadian) urban areas. This section considers urban areas throughout the globe. Three databases are evaluated. The first, the Global Rural–Urban Mapping Project (GRUMP), provides a comprehensive cross-sectional dataset of population size and land area for 55384 urban areas in 218 countries (CIESIN, 2004). GRUMP values are currently an ‘alpha’ version—i.e. still in development (CIESIN, 2004); thus, values presented next should be considered preliminary. Because the GRUMP data are cross-sectional rather than panel or time-series, they cannot generate estimates for the parameter n. Rather, the GRUMP data are used here to estimate LPD based on the average population and land area for all urban areas within each country.

Median and population-weighted mean population density values in GRUMP are ~50 per cent higher than values for US Census urban areas. Compared with US urban areas, GRUMP data are similar for median LPD and ~3 times lower for population-weighted mean LPD.

The second database provides urban population and land area for 394 cities in 17 countries (Table 8 in United Nations, 2000a). Figure 6 shows general similarities between trends in population density and in LPD for three countries presented (Thailand, India and the US; these three countries account for 77 per cent of cities in the database). Two attributes of Figure 6 are

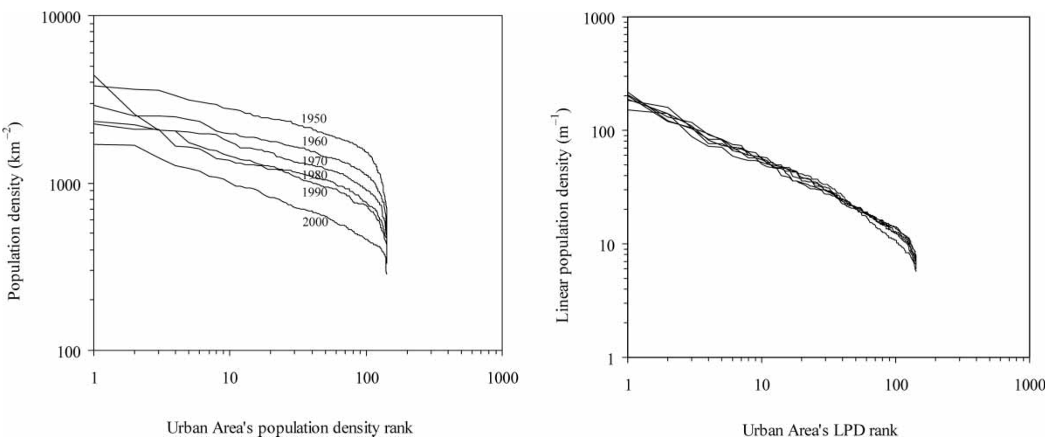

Figure 5. Relationship between population density and population density rank (left) and between LPD and LPD rank (right). Values plotted are for Group 1 urban areas, for the six censuses during 1950–2000. This figure mirrors a similar plot used to investigate Zipf’s (1949) population rank–size rule (Pumain, 2003). The distributions shown here support a generalised version of Zipf’s rule for population density and for LPD.

Table 2. Modified rank–size rule parameters for Group 1 urban areas

	Largest value in dataset[a]		K		α		R^2	
	1950	2000	1950	2000	1950	2000	1950	2000
Population	12 million	18 million	15 million	57 million	−1.06	−1.20	0.99	0.93
Urban land area	3,200 km^2	14,000 km^2	5,900 km^2	50,000 km^2	−0.97	−1.02	0.96	0.88
LPD	220 m^{-1}	150 m^{-1}	230 m^{-1}	280 m^{-1}	−0.62	−0.71	0.98	0.95
Population density	3,800 km^{-2}	1,700 km^{-2}	6,200 km^{-2}	2,300 km^{-2}	−0.32	−0.35	0.74	0.95

[a]Largest dataset values are presented here for comparison; they are not a parameter in the modified rank–size rule.

mathematically necessary results of LPD being relatively more dependent on population compared with population density (i.e. density is PA^{-1} and LPD is $PA^{-0.5}$; thus, population variability among urban areas yields greater variability in LPD than in density): first, urban population is a better predictor of LPD than of population density (i.e. R^2 values for the best-fit regression lines are greater in the right plot than in the left plot); and, secondly, in any specific urban population, there is less within-country and between-country variability for LPD than for population density (i.e. when comparing data points for one country or when comparing among the three best-fit lines, for a given urban population there is a narrower range of y-axis values in the right plot than in the left plot).

The third database, a panel dataset by Kenworthy and Laube (2000), contains population and land area by decade during 1960–1990 for 35 urban areas. Analyses here consider the 26 areas with a greater than 10 per cent increase in population during 1960–1990. These areas in the Kenworthy and Laube (2000) dataset have relatively large populations and population densities: the year 1990 median population is 2.5 million (22 times larger than in the year 1990 US census dataset) and the year 1990 median population density is 1550 km^{-2} (twice as large as in the US). Values for the parameter n (mean = 1.7; median = 1.5; range: 0.6–3.9)

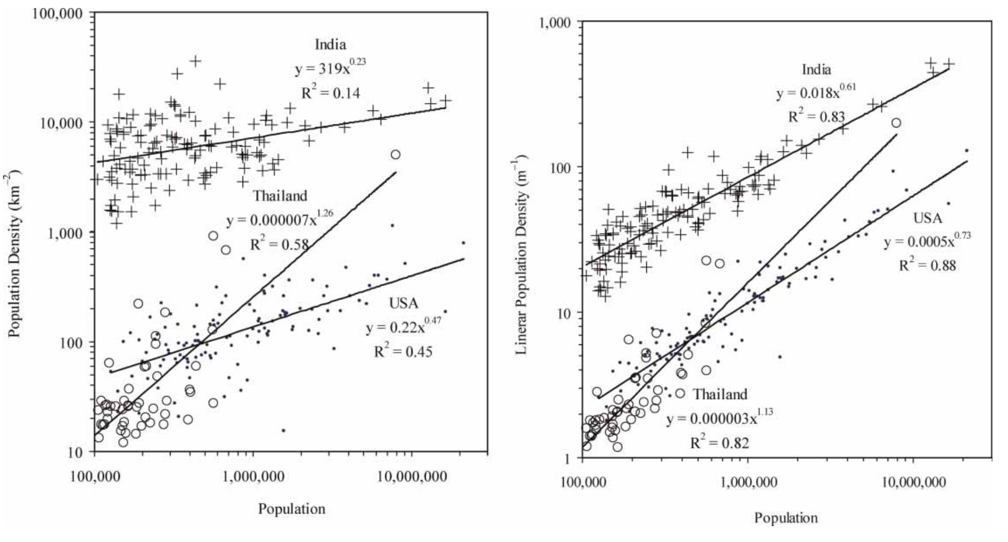

Figure 6. Relationship of urban population to population density (left) and to LPD (right) for India, Thailand, and the US. Differences among the three countries are similar in the left and right plots. Urban population is a better predictor of LPD than population density (i.e. R^2 values for the best-fit line are larger in the right plot than in the left plot). Data are from the United Nations (2000a).

are similar to, or slightly smaller than, values reported above for the US and Canada. Median LPD values for the Kenworthy and Laube (2000) data are roughly constant over time and are significantly larger than for the US (less than 5 per cent change during 1960–1990; value: $62\,m^{-1}$, which is six times higher than for the US). Mean and population-weighted mean LPD values increased during 1960–1990, but the changes (from $82\,m^{-1}$ to $96\,m^{-1}$ for mean; from $175\,m^{-1}$ to $208\,m^{-1}$ for population-weighted mean) are less than 20 per cent. Cumulative distributions plots for LPD values (not shown) in 1960 and in 1990 are similar to each other. Thus, although this international panel dataset is relatively small and contains comparatively large urban areas, it corroborates the observation from the US census data that the shape and magnitude of the LPD distribution, including mean and median values, are approximately constant over multiple decades.

Implications

LPD is worth careful consideration mainly because of long-term stability in the distribution of values. LPD appears to conform well to a size–rank rule. Even though consistency in average LPD is suggested by the value $n = 2$ for equation (1), there is sufficient variability in estimated values for n among UAs and among area groups that greater variability in average LPD might be expected. Average LPD has changed less than 10 per cent over multiple decades for the UAs studied here. Few, if any, other average attributes of urban form have shown such consistency over the past half-century. Based on this long-term consistency, LPD may be useful for corroborating and calibrating urban growth models, especially those considering both within- and between-city migration.

Additional work is necessary to explain these findings, relate them to within-city urban dynamics (such as mobility patterns, jobs/housing balance, polycentric urban form) and to understand their implications. Many factors influence expansion of the urban land area, including physical geography; transport systems and other infrastructure; economic factors such as supply and demand for housing, employment and quality education; prevalence of amenities (such as parks, restaurants) and disamenities (such as crime); and, social factors such as norms, values and culture. A similarly long list of factors is influenced by expansion of the urban land area.

Future research may connect the observations in this work with individuals' decisions and environmental interactions (Anas *et al.*, 1998). For example, to speculate, perhaps modern vehicle-centric urban form leads people to interact with their environment in a linear manner, such as along transport corridors, rather than in terms of the two-dimensional space around them, and this fact leads to long-term consistency in linear population density. Or, perhaps the fact that each urban newcomer, on average, occupies about twice the per capita land area of existing residents reflects people's desire to consume more land than their peers, while not exceeding income constraints and/or appearing overextravagant. Economists argue that the free-market equilibrium size of an urban area reflects the slope of the bid-rent curve and the cost of developing new land at the urban boundary. From this standpoint, results presented here reflect multidecadal changes among cities in income, housing costs and transport costs.

The analyses presented here are empirical—a series of observations about urban growth—rather than normative. For example, while LPD distributions are observed to be consistent over time, the article does not discuss whether planners should try to shift this pattern. Given the significant debate in planning literature and practice about whether and how to influence population density, it may be worthwhile to apply normative questions to LPD if this metric becomes more widespread.

One practical application of findings presented here relates to health effects of motor vehicle emissions. The environmental health impact attributable to an emission source, such as motor vehicles in an urban area, can often be estimated as the product of three terms: emission rate (mass per day), intake fraction (mass inhaled per mass emitted) and toxicity (health impact per mass inhaled) (Bennett *et al.*, 2002). Intake fraction, a dimensionless number ranging from zero to one, quantifies the 'exposure efficiency' of an urban area. For example, an intake fraction of 15 per million, which is a typical value for certain types of vehicle emission in US urban areas (Marshall *et al.*, 2005b), would mean that, on average, 15 mg are inhaled collectively by urban residents per kg emitted. A larger intake fraction value would indicate that vehicle emissions are more efficiently delivered to people's lungs—something that is to be avoided (reduced) when possible.

Intake fraction depends on several factors, including the size of the exposed population, proximity between emissions and people, dilution rate of emissions owing to atmospheric mixing and persistence of a pollutant in the environment. Intake fraction for non-reactive atmospheric vehicle emissions (such as carbon monoxide and benzene) in urban areas is proportional to LPD, not population density (Lai *et al.*, 2000; Marshall *et al.*, 2005b). (The same is likely to hold for sources similar to motor vehicles—for example, outdoor 'area sources' that are spatially well-distributed, such as gasoline lawn mowers and dry cleaners.) Results presented here suggested that average proximity between people and vehicles in US urban areas, as measured by average LPD, has not changed significantly over the past several decades. Changes in population-weighted LPD are modest (for example, a 25 per cent decline during 1950–2000 for Group 1). Thus, to a first approximation, intake fraction of urban vehicle emissions has remained approximately constant over multidecadal time-scales.

Urban planning scenarios are often used to forecast urban motor vehicle emissions. If the intake fraction of vehicle emissions in a given urban area were constant over time, then the exposure impact per mass of vehicle emission would also be constant. (Of course, to understand the environmental health impacts of vehicles, one would also need to track over time the two other terms besides intake fraction—namely, emissions and toxicity. For most or all pollutants, toxicity is assumed constant over time, owing in part to a lack of evidence to the contrary.) Based on results presented here, intake fraction may increase or decrease in any given area, but US mean and median values appear roughly constant over time. Because intake fraction varies significantly among urban areas (and between urban and rural areas), the exposure impact per mass of vehicle emission varies significantly among urban areas (and between urban and rural areas). Findings presented here are important for urban planners, environmental health practitioners and air pollution exposure and risk assessors.

The findings suggest that intake fraction varies more among urban areas than population density does. For example, Los Angeles is 3.5 times denser than the median UA (1700 versus 490 km^{-2}), but LPD—and hence, to a first approximation, intake fraction—is 18 times greater for Los Angeles than the median UA (141 versus 7.8 m^{-1}). Urban areas with higher intake fraction values might wish to seek more stringent emission controls.

Finally, if vehicle usage and levels of walking and other exercise are proportional to population density (Ewing and Cervero, 2001; Frank *et al.*, 2005; Frumpkin, 2002; Holtzclaw *et al.*, 2002) but intake fraction is proportional to LPD, this would suggest that small-population, high-density urban areas such as Ames, IA—which has a large population density but a small LPD—may be a planning goal from the standpoint of vehicle energy consumption, air pollution and public health (see Marshall *et al.*, 2005a). Similar objectives—reduced vehicle energy consumption, increased exercise levels and reduced proximity between people and vehicle emissions—may also be achievable via neighbourhood-scale urban design choices.

Conclusion

Data for US urban areas (UAs) during 1950–2000 were analysed to investigate changes over time in urban land area size (A) and population size (P). The scaling relationship $A \propto P^n$, where n is an empirical constant, appears to hold reasonably well when population size increases by more than 10 per cent. Typical values for n are 1.3–3.7, with a median value of $\sim$2. If $n = 2$, the parameter group $PA^{-0.5}$ ('linear population density') would be approximately constant over time. Linear population density (LPD) distributions for the UA groups investigated are, in fact, approximately constant during 1950–2000. Few, if any, other average attributes of urban form are so consistent over time. Results provide a straightforward mathematical description of long-term population mobility within and among urban areas.

In general, when comparing UAs, LPD varies more than population density; but, when comparing over time (either for a single urban area or for averages among areas), population density varies more than LPD. Population density and LPD conform well to generalised forms of Zipf's (1949) rule. Two international datasets were investigated: one with cross-sectional data and one with panel data. The panel dataset, although small (26 urban areas with moderate or rapid growth), also reveals multidecadal consistency in LPD distributions. Intake fraction of urban vehicle emissions is one practical application of the findings presented here.

Additional work could usefully confirm these findings for other datasets (for example, Angel *et al.*, 2005 and Robinson *et al.*, 2000), investigate the theoretical underpinnings and explore further the practical implications.

References

Anas, A., Arnott, R. and Small, K. A. (1998) Urban spatial structure, *Journal of Economic Literature*, 36, pp. 1426–1464.

Anas, A., Arnott, R. J. and Small, K. A. (2000) The panexponential monocentric model, *Journal of Urban Economics*, 47, pp. 165–179.

Anderson, J. E. (1985) The changing structure of a city: temporal changes in cubic-spline urban density patterns, *Journal of Regional Science*, 25, pp. 413–425.

Angel, S., Sheppard, S. C. and Civco, D. L. (2005) *The Dynamics of Global Urban Expansion*. Washington, DC: Transport and Urban Development Department, The World Bank (http://www.williams.edu/Economics/UrbanGrowth/WorkingPapers.htm).

Batty, M. and Kim, K. S. (1992) Form follows function: reformulating urban population density functions, *Urban Studies*, 29, pp. 1043–1069.

Bennett, D. H., McKone, T. E., Evans, J. S. et al. (2002) Defining intake fraction, *Environmental Science and Technology*, 36, pp. 206A–211A.

Bettencourt, L. M. A., Lobo, J., Helbing, D. et al. (2007) Growth, innovation, scaling, and the pace of life in cities, *Proceedings of the National Academy of Sciences*, 104, pp. 7301–7306.

Black, D. and Henderson, V. (2003) Urban evolution in the USA, *Journal of Economic Geography*, 3, pp. 343–372.

Bunting, T., Filion, P. and Priston, H. (2002) Density gradients in Canadian metropolitan regions, 1971–96: differential patterns of central area and suburban growth and change, *Urban Studies*, 39, pp. 2531–2552.

CIESIN (Center for International Earth Science Information Network) (2004) Global rural–urban mapping project (GRUMP): urban extents, Version 1 alpha. CIESIN, Columbia University, New York (http://beta.sedac.ciesin.columbia.edu/gpw/global.jsp).

Clark, C. (1951) Urban population densities, *Journal of the Royal Statistical Society*, 114, pp. 490–496.

CNSTAT (Committee on National Statistics) (2004) The 2000 coverage evaluation program, in: C. F. Citro, D. L. Cork and J. L. Norwood (Eds) *The 2000 Census: Counting under Adversity*, pp. 185–268. Washington, DC: The National Academies Press (http://www.nap.edu/catalog/10907.html).

Edmonston, B., Goldberg, M. A. and Mercer, J. (1985) Urban form in Canada and the United States: an examination of urban density gradients, *Urban Studies*, 22, pp. 209–217.

Ewing, R. and Cervero, R. (2001) Travel and the built environment: a synthesis, *Transportation Research Record*, 1780, pp. 87–113.

Frank, L. D., Schmid, T. L., Sallis, J. F. et al. (2005) Linking objectively measured physical activity with objectively measured urban form, *American Journal of Preventative Medicine*, 28, pp. 117–125.

Frumpkin, H. (2002) Urban sprawl and public health, *Public Health Reports*, 117, pp. 201–217.

Gabaix, X. (1999) Zipf's law for cities: an explanation, *Quarterly Journal of Economics*, 113, pp. 739–767.

Hogan, H. and Robinson, G. (2000) *What the Census Bureau's coverage evaluation programs tell us about differential undercount*. Paper presented at the *Conference on Undercounted Ethnic Populations*, Richmond, Virginia, May 1993 (http://www.census.gov/population/www/documentation/1993/conference.html).

Holtzclaw, J., Clear, R., Dittmar, H. *et al.* (2002) Location efficiency: neighborhood and socioeconomic characteristics determine auto ownership and use—studies in Chicago, Los Angeles and San Francisco, *Transportation Planning and Technology*, 25, pp. 1–27.

Ioannides, Y. M. and Overman, H. G. (2003) Zipf's law for cities: an empirical examination, *Regional Science and Urban Economics*, 33, pp. 127–137.

Kenworthy, J. R. and Laube, F. B. (2000) *An International Sourcebook of Automobile Dependence in Cities, 1960–1990*. Boulder, CO: University Press of Colorado.

Lai, A. C. K., Thatcher, T. L. and Nazaroff, W. W. (2000) Inhalation transfer factors for air pollution health risk assessment, *Journal of the Air and Waste Management Association*, 50, pp. 1688–1699.

Marshall, J. D., McKone, T. E., Deakin, E. and Nazaroff, W.W. (2005a) Inhalation of motor vehicle emissions: effects of urban population and land area, *Atmospheric Environment*, 39, pp. 283–295.

Marshall, J. D., Teoh, S.-K. and Nazaroff, W. W. (2005b) Intake fraction of nonreactive vehicle emissions in US urban areas, *Atmospheric Environment*, 39, pp. 1363–1371.

Mills, E. S. and Tan, J. P. (1980) A comparison of urban population density functions in developed and developing countries, *Urban Studies*, 17, pp. 313–321.

Mulry, M. (2006) *Summary of accuracy and coverage evaluation for Census 2000*. Research Report Series: Statistics No. 2006–3, US Department of Commerce, Bureau of the Census, Washington, DC (http://www.census.gov/srd/papers/pdf/rrs2006–03.pdf).

Murakami, A., Zain, A. M., Takeuchi, K. *et al.* (2005) Trends in urbanization and patterns of land use in the Asian mega cities Jakarta, Bangkok, and Metro Manila, *Landscape and Urban Planning*, 70, pp. 251–259.

Newling, B. E. (1966) Urban growth and spatial structure: mathematical models and empirical evidence, *The Geographical Review*, 59, pp. 242–252.

Pumain, D. (2003) *Scaling laws and urban systems*. Working Paper No. 04–02–002, Santa Fe Institute (http://www.santafe.edu/research/publications/workingpapers/04-02-002.pdf).

Robinson, J. A., McRay, B. and Lulla, K. P. (2000) Twenty-eight years of urban growth in North America quantified by analysis of photographs from Apollo, Skylab and Shuttle-Mir, in: K. P. Lulla and L. V. Dessinov (Eds) *Dynamic Earth Environments: Remote Sensing Observations from Shuttle-Mir Missions*, pp. 25–41, 262, 269–270. New York: John Wiley & Sons.

Robinson, J. G., Ahmed, B., Gupta, P. D. and Woodrow, K. A. (1993) Estimation of population coverage in the 1990 United States Census based on demographic analysis, *Journal of the American Statistical Association*, 88, pp. 1061–1071.

Rosser, J. B. (1980) The dynamics of ghetto boundary movement and ghetto shape, *Urban Studies*, 17, pp. 231–235.

Skaburskis, A. (1989) Inversions in urban density gradients: a brief look at the Vancouver metropolitan area's density profile, *Urban Studies*, 26, pp. 397–401.

Stern, D. I. (1993) Historical path-dependence of the urban population density gradient, *The Annals of Regional Science (Historical Archive)*, 27, pp. 259–283.

United Nations (2000a) *Demographic Yearbook*. New York: United Nations Statistics Division (http://unstats.un.org/unsd/demographic/products/dyb/dyb2000.htm).

United Nations (2000b) *World Urbanization Prospects: The 1999 Revision*, (www.un/org/esa/population/pubsarchive/urbanization/urbanization.pdf).

Urzua, C. M. (2000) A simple and efficient test for Zipf's law, *Economics Letters*, 66, pp. 257–260.

US Census (1992) *1990 Census of population, general population characteristics, urbanized areas, section 3 of 3*. Report 1990–CP–1–1C, US Department of Commerce, Bureau of the Census, Washington, DC (http://www.census.gov/population/www/censusdata/90pubs/cp-1.html).

US Census (2004) *Census 2000 urban and rural classification*. US Department of Commerce, Bureau of the Census, Washington, DC (http://www.census.gov/geo/www/ua/ua_2k.html).

Zheng, X.-P. (1991) Metropolitan spatial structure and its determinants: a case-study of Tokyo, *Urban Studies*, 28, pp. 87–104.

Zipf, G. K. (1949) *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wiley.

Appendix

Equations to estimate the parameter n in equation (1), employing the exponential population density gradient (Clark, 1951), are presented in Table A1. Parameter abbreviations are as follows: $\rho(r)$ is the population density (km^{-2}), which is a function of distance (r; km) from the centre of the urban area; ρ_0 is the population density (km^{-2}) at the centre (at $r = 0$); λ is the exponential decay coefficient (km^{-1}); $\tilde{\rho}$ is the threshold density separating 'urban' from 'not urban'—the value employed in this work, 1000 mile^{-2} (346 km^{-2}), is from the US census; $\tilde{r}$ is the distance (km) from the centre to the edge of the urban area; A is the urban land area (km); P is the population size; and, ρ_{ave} is the average population density (km^{-2}). The subscripts '1' and '2' (A_1, A_2, etc.) refer to times 1 and 2.

Parameter groups used in the table are

$$\beta \equiv \ln\left(\frac{\rho_o}{\tilde{\rho}}\right) \text{ and}$$

$$\Phi \equiv [1 - e^{-\tilde{r}\lambda}(\tilde{r}\lambda + 1] = \left[1 - \frac{\tilde{\rho}}{\rho_o}(\beta + 1)\right]$$

For US urban areas, typical values for β and Φ are in the ranges 2–4 and 0.65–0.90 respectively. These two groups (β and Φ) are employed because their variability among urban areas is small relative to variability in parameters such as population and area.

Table A1. Estimating the parameter n in equation 1, employing the exponential population density gradient by Clark (1951)

Parameter for the urban area	Equation
Population density gradient (Clark, 1951)	$\rho(r) = \rho_o e^{-r\lambda}$
Maximum and minimum population density	ρ_o and $\tilde{\rho}$ respectively
Radius	$\tilde{r} = \frac{\beta}{\lambda}$
Area	$A = \pi\tilde{r}^2 = \frac{\pi\beta^2}{\lambda^2}$
Population	$P = \int_0^{\tilde{r}} 2\pi r\rho(r)dr = \frac{2\pi\rho_o\Phi}{\lambda^2}$
Average population density	$\rho_{ave} = \frac{P}{A} = \frac{2\rho_o\Phi}{\beta^2}$
The parameter n	$n = \frac{\ln(A_2/A_1)}{\ln(P_2/P_1)} = \frac{2\ln(\beta_2\lambda_1/\beta_1\lambda_2)}{\ln(\rho_{o,2}\Phi_2/\rho_{o,1}\Phi_1)+2\ln(\lambda_1\lambda_2)}$.