INTERFACE

Research

Check for updates

Authors for correspondence:
Vicky Chuqiao Yang
e-mail: vcy@santafe.edu
Christopher P. Kempes
e-mail: ckempes@santafe.edu

THE ROYAL SOCIETY
PUBLISHING

Scaling of urban income inequality in the USA

Elisa Heinrich Mora[1,2], Cate Heine[2,3], Jacob J. Jackson[2,4], Geoffrey B. West[2], Vicky Chuqiao Yang[2] and Christopher P. Kempes[2]

[1]Minerva Schools at KGI, San Francisco, CA 94103, USA
[2]Santa Fe institute, Santa Fe, NM 87501, USA
[3]Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Brown University, Providence, RI 02912, USA

EHM, 0000-0003-1204-2526; VCY, 0000-0001-6111-3699

Urban scaling analysis, the study of how aggregated urban features vary with the population of an urban area, provides a promising framework for discovering commonalities across cities and uncovering dynamics shared by cities across time and space. Here, we use the urban scaling framework to study an important, but under-explored feature in this community—income inequality. We propose a new method to study the scaling of income distributions by analysing total income scaling in population percentiles. We show that income in the least wealthy decile (10%) scales close to linearly with city population, while income in the most wealthy decile scale with a significantly superlinear exponent. In contrast to the superlinear scaling of total income with city population, this decile scaling illustrates that the benefits of larger cities are increasingly unequally distributed. For the poorest income deciles, cities have no positive effect over the null expectation of a linear increase. We repeat our analysis after adjusting income by housing cost, and find similar results. We then further analyse the shapes of income distributions. First, we find that mean, variance, skewness and kurtosis of income distributions all increase with city size. Second, the Kullback–Leibler divergence between a city's income distribution and that of the largest city decreases with city population, suggesting the overall shape of income distribution shifts with city population. As most urban scaling theories consider densifying interactions within cities as the fundamental process leading to the superlinear increase of many features, our results suggest this effect is only seen in the upper deciles of the cities. Our finding encourages future work to consider heterogeneous models of interactions to form a more coherent understanding of urban scaling.

1. Introduction

Throughout human history, the global urban population has grown continuously. More than half of the global population is currently urbanized, placing cities at the centre of human development [1]. It is estimated that by 2030, the number of megacities, cities with more than 10 million inhabitants, will increase from 10 to approximately 40 [1]. Thus, there is an urgent need for a quantitative and predictive theory for how larger urban areas affect a wide variety of city features, dynamics and outcomes [2,3]. Perhaps most critically, we need this theory to address how larger cities positively and negatively affect socioeconomic outcomes and the quality of life of individuals.

Previous research has demonstrated power-law-like relationships between urban population (also referred to as size later in the text) and many urban features such as GDP, patents, crime and contagious diseases that persist globally [4–8]. These relationships can often be described by

$$Y = Y_0 N^\beta, \tag{1.1}$$

where Y is an urban feature, such as GDP or number of crime instances, N is the population of the city, Y_0 is a constant and β is the scaling exponent. For many

urban outputs, the scaling exponent β is greater than 1, suggesting greater rates of productivity (in both the positive and negative sense) in more populated cities. These observations, known as urban scaling, suggest that a small set of mechanisms significantly influence a variety of urban features across diverse cities [9,10]. Understanding these mechanisms has important implications for developing more prosperous and safer cities. In this framework, desirable aspects with $\beta > 1$ have positive returns to scale, while desirable aspects with $\beta < 1$ have a less than linear return to scale, demonstrating a diseconomy of scale. Similarly, for undesirable features $\beta > 1$ shows a diseconomy of scale since the associated *per capita* costs would be increasing with city size.

One important aspect of urban features that remains under-explored in the urban scaling framework is economic inequality. Inequality has fundamental implications for individuals' quality of life and the productivity and stability of societies [11]. Past research has heightened debate about economic inequality and its relationship with economic growth and general welfare [12–18]. Many have raised concern of its negative effects on political stability [12,19], crime [20] and corruption [21]. It has been shown that more unequal places have higher murder rates, grow more slowly, and the correlation between area-level inequality and population growth is positive [22]. Economic inequality is usually measured in terms of the dispersion in the distribution of income or wealth, such as in the Gini coefficient. Some past research has noted larger cities are correlated with increasing Gini coefficient in income distribution [23–25], but it remains unclear if there are systematic relationships between other features of the income distributions and urban area size. Furthermore, characterizing distributions by a single metric may lose important information [12]—for example, does being poor in bigger cities correspond to a higher or lower standard of living than being poor in a smaller city?

A few recent studies [26,27] have investigated the scaling of total income in various income brackets in Australia. These studies find that the total income in lower income brackets scales sublinearly or linearly, while higher brackets scale superlinearly, suggesting greater income agglomeration in the higher income categories in more populated cities. While these studies are informative and provide a new measure for inequality in terms of absolute income (instead of relative income, as in the Gini coefficient), a limitation is that this measure confounds inequality with average income, which increases with city population. In particular, the 'equal' situation in this new measure of inequality is when the total income for all income brackets scales linearly. However, given that total income scales superlinearly in cities globally [4,5], this 'equal' situation is unlikely to occur. For example, even if the shapes of income distributions remain identical, income bracket aggregations follow distinct scaling relationships as a result of differences in mean. Figure 1*a,c* illustrates this behaviour using simulated log-normal distributions. While the measure of inequality proposed in [26,27] can be valuable for some applications, it would be useful to untangle the increase in mean from the greater dispersion in income.

In this manuscript, we address a few keys questions: (1) how does income inequality (adjusted for shifting average income) systematically change with city size? (2) How different is the income of rich and poor people (measured by percentiles of the population) in small and large cities, and how does this difference scale with city size? (3) Are poor people in a larger city better off than poor people in a small city, after adjusting by the cost of living? How about the same for rich people?

Here, we propose a new method to study the scaling of inequality by analysing total income scaling in population percentiles. We show that income in the least wealthy decile (10%) scales almost linearly with city size, while that in the most wealthy decile scales with a significantly superlinear exponent. This illustrates that the benefits of larger cities are increasingly unequally distributed, and for the poorest income deciles, city growth has no positive effect on income growth over the null expectation of a linear increase. We find that these results hold after adjusting for cost of living as proxied by housing cost. We then introduce systematic considerations of the entire distribution of income to show which income distribution features are changing with city size. We find that the mean, variance, skewness and kurtosis of the income distribution all scale systematically with city size. We introduce a KL-divergence procedure to systematically compare all moments and find that comparisons with the largest cities also demonstrate a systematic scaling with city size, indicating that the overall shape of income distribution is radically shifting with city size. Finally, we discuss how these observations can be connected with the proposed mechanisms underlying urban scaling.

2. Data and methods

2.1. Data and income distribution estimation

The primary dataset used in our analysis is the 2015 American Community Survey conducted by the US Census Bureau (see electronic supplementary material for more detail). We use the income data reported on the level of census tracts, small local areas of on average 4500 people, of which on average 2300 reported income. In order to aggregate census tracts into cities, we gather the geographical definitions of urban areas from the United States Office of Management and Budget, which include Metropolitan Statistical Areas and Micropolitan Statistical Areas. We perform a spatial join between urban area outlines and census tract outlines. All census tracts which intersect a given urban area are assigned to that urban area. Census tracts which intersect multiple urban areas are assigned to both. The population of an urban area is defined as the sum of their census tracts' populations. In our study sample, we analyse urban areas (also referred to as cities) with a population greater than 100 000.

We infer the individual-level income distribution in cities by applying the Gaussian kernel density estimator with a widened Silverman bandwidth function on the census-tract-level data. This method assumes income in each census tract is distributed as a Gaussian. The mean equals the average income of the census tract, and the standard deviation is calculated as a function of the number of data points. Aggregating the Gaussian probability density functions (PDFs) for each census tract in the city produces an estimated income PDF for the city. Examples of the estimated individual-level income distribution for a few cities are shown in figure 2.
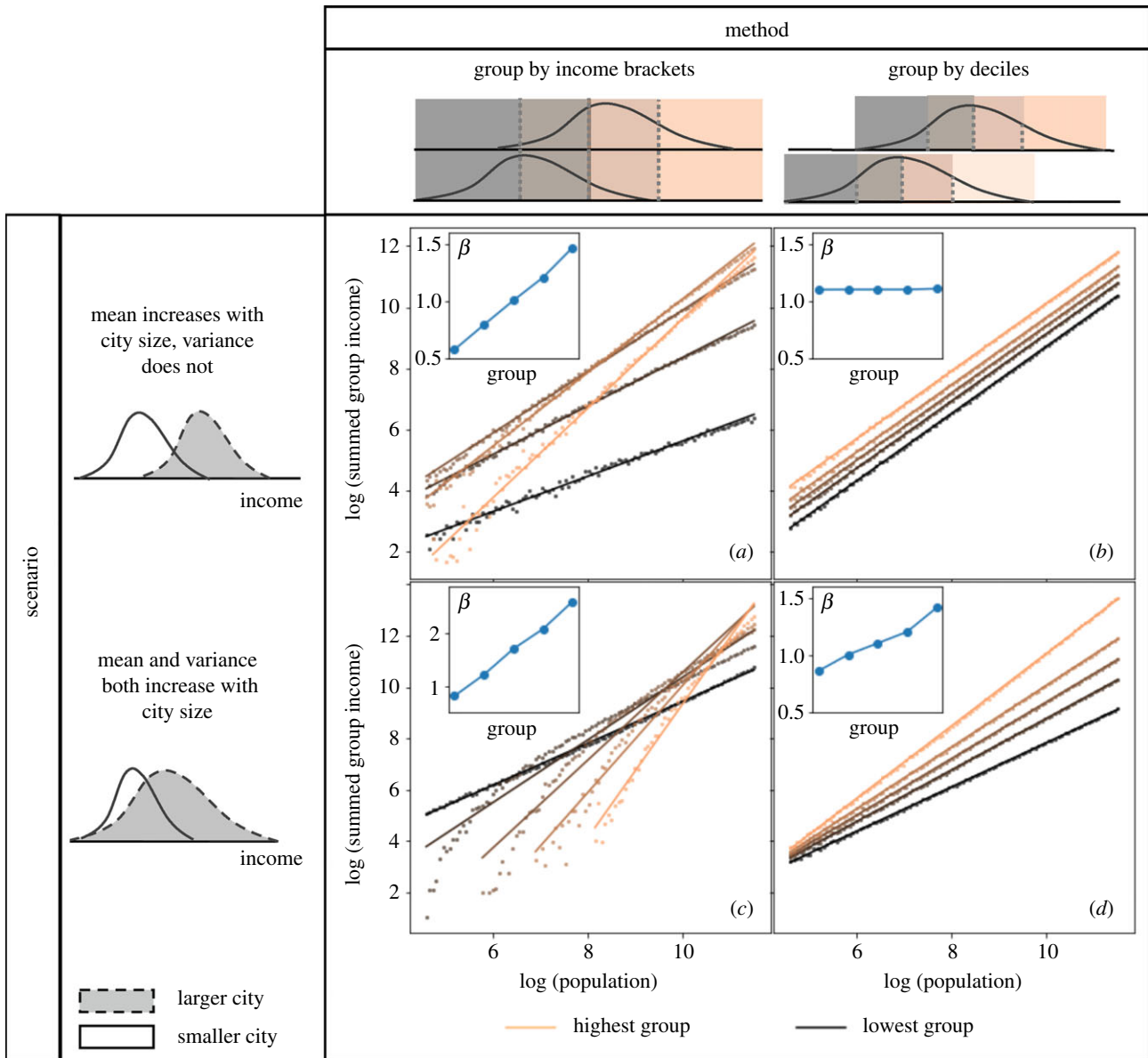
Figure 1. Illustration comparing two methodologies—scaling obtained from grouping by income bracket (*a,c*) and that by decile (*b,d*). Using simulated log-normal income distributions in two scenarios—log-mean increases with city size while log-variance remains the same (*a,b*), and log-mean and log-variance both increases with city size (*c,d*). The income distributions are illustrated on a log-scale. The income-bracket grouping (*a,c*) leads to differences in the groups' income scaling for both scenarios, and fails to distinguish whether larger cities have more dispersion in their income distributions. The decile grouping (*b,d*) leads to the differences in the groups' scaling observed only when the dispersion increases with the population. The insets show how scaling exponent (*β*) varies with income groups (bracket or decile).

2.2. Analysis of income scaling in deciles

We propose a new method to investigate the scaling of income aggregated by deciles in each city (i.e. the bottom 10%, the next 10% and so on). The number of individuals in decile n of city i is, $N_i^{(n)} = N_i/10$, where N_i is the population of city i.

The total income in decile n of city i, $Y_i^{(n)}$ is,

$$Y_i^{(n)} = \sum_{j \in \mathbf{D}^{(n)}} y_{i,j}, \qquad (2.1)$$

where $\mathbf{D}^{(n)}$ are the individuals in income decile n, and $y_{i,j}$ is the income of individual j in city i. See electronic supplementary material for more details on the decile assignment in our computational implementation.

Figure 1*c,d* illustrates this method on simulated log-normal income distributions. Panel *c* represents the situation

in which cities shift in log-mean with city size, but do not shift in log-standard deviation, and panel *d* represents the situation in which cities increase both log-mean and log-standard deviations with city size. We consider the former case an example of the 'equal' situation, and this method should lead to no variation in scaling exponents across deciles. Variations in scaling exponent only occur for the latter case. We also contrast the results of our method with that of the grouping by income bracket method in figure 1*a,c*, where variations in scaling exponents occur for both scenarios.

We group the data of income distributions of each city into deciles: the 10% of the population which reports the lowest income is grouped into the first decile (decile 1), and likewise for all 10 deciles up to the 10% of the population which has the highest income (decile 10). We then estimate the scaling exponent of total income for all deciles. We estimate the scaling exponent, $β$, and corresponding confidence intervals,
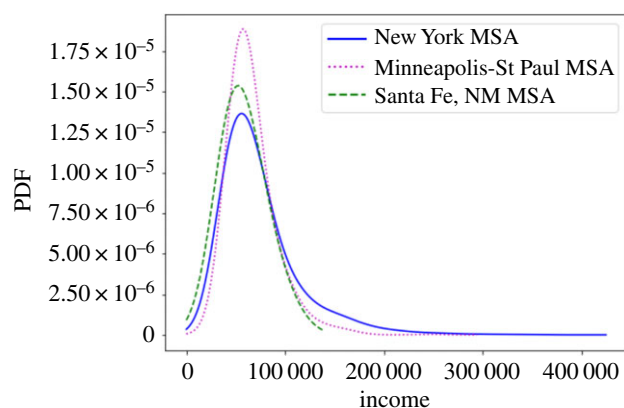
Figure 2. Examples of the estimated income distributions using census tract data. Income is measured in US dollars. The three metropolitan areas shown are: New York–Newark–Jersey City, NY–NJ–PA, population 20 316 622; Minneapolis-St Paul-Bloomington, MN-WI, population 3 670 397; Santa Fe, NM, population 204 396.

by performing an ordinary least square regression of the log-transformed variables, $\log(Y_i^{(n)}) = \beta \log(N_i^{(n)}) + c$, and β and c are the fitted parameters. This methodology is consistent with previous research such as [4].

2.3. Adjusting income by housing cost

In order to normalize income by the cost of living, we calculate total housing cost in a census tract as $\text{cost} = 12\,(u_{\text{rent}}\,r + u_{\text{own}}\,o)$, where the average monthly rent r, the average monthly owner costs o, and the number of units of each type u_{rent} and u_{own} are all taken from the 2015 American Community Survey (see electronic supplementary material for more detail and access information). We then repeat the decile-grouped analysis on income adjusted for housing cost, as well as analyse how the proportion of income spent on housing varies with city size in each decile.

2.4. Analysis of distributions

We further analyse how the shapes of the income distributions vary with city population. We first compute the first four statistical moments, mean, variance, skewness and kurtosis, for income distributions of each city, and analyse how they vary with population. We then compute the Kullback–Leibler (KL) divergence between each city's income distribution and that of the largest city (New York–Newark–Jersey City area). The KL divergence measures how different one distribution is from another, while the zero value indicates the two distributions are identical, and a greater value indicates more divergence. Mathematically, the KL divergence between two discrete distributions of random variable x, $P(x)$ and $Q(x)$ is,

$$KL(P\|Q) = \sum_x P(x) \log\left(\frac{P(x)}{Q(x)}\right). \tag{2.2}$$

3. Results

3.1. Scaling of income in deciles

The results for scaling of income aggregated in deciles are summarized in figure 3. For the lowest two deciles, the scaling exponent β is linear or slightly sublinear (0.97). For upper

deciles, β is consistently superlinear, as high as 1.16 when compared with the scaling exponent of total income in our dataset, $\beta = 1.07$. This shows that scaling effects are not equivalent for all segments of the population. The poorest two deciles in bigger cities make about the same income as their counterparts in smaller cities, while the wealthiest eight deciles in bigger cities make more than their counterparts in smaller cities, where the difference increases with the decile.

3.2. Scaling of decile income adjusted by housing cost

While the differences in income scaling that we have identified are important, they are not necessarily grounded in differences in the experiences of urban residents—cost of living can vary drastically across and within US cities, and if cost of living is changing in the exact same way as income, differences in income scaling between groups begin to lose meaning. In order to understand whether the differences in income scaling we see between deciles create differences in affordability and purchasing power, we look at changes in housing cost with city size.

We find that aggregate housing cost scales faster than aggregate income for *every* decile, implying that while income per person increases with city size, larger cities may still be overall less affordable. This difference is more dramatic for the poorer deciles—in the bottom decile, housing cost scales with $\beta = 1.11$ while income scales with $\beta = 1.01$; in the top decile, housing cost scales with $\beta = 1.29$ while income scales with $\beta = 1.27$. This is visualized in figure 4*a*—income exponents begin to catch up to housing cost exponents in richer deciles, but income exponents are never as high as housing cost exponents. Perhaps more intuitively, in figure 4*b*, we can see that the ratio between total housing cost and total income grows with city size for every decile, but more dramatically for poorer deciles. Together, these results imply a widening gap between richer and poorer residents in affordability of cities with city size.

3.3. Analysis of income distribution characteristics

We further analyse how income distributions vary with urban area population by studying the statistical moments of the income distributions. We first examine the first four moments: mean, variance, skewness and kurtosis.

The scaling of the four moments of the estimated individual income distribution for all cities in our data is shown in figure 5. The first moment, the mean, shows the well-characterized urban agglomeration effect: *per capita* income increases with city size [4]. The second, and third moments both increase similarly with city population, suggesting a widening of the distribution and increasing asymmetry with greater urban population. This can also be qualitatively observed in the example distributions in figure 2. Lastly, the kurtosis also increases with population size, showing an increasingly heavy tail with greater urban population.

We find a stronger relationship for higher statistical moments, indicating that for larger American cities, there is a more evident increase in the third and fourth moments. This means that there is a stronger increase in the growing tail of the distribution, in comparison to the first two statistical moments. This gives us an interesting indication of the distribution of economic benefits.

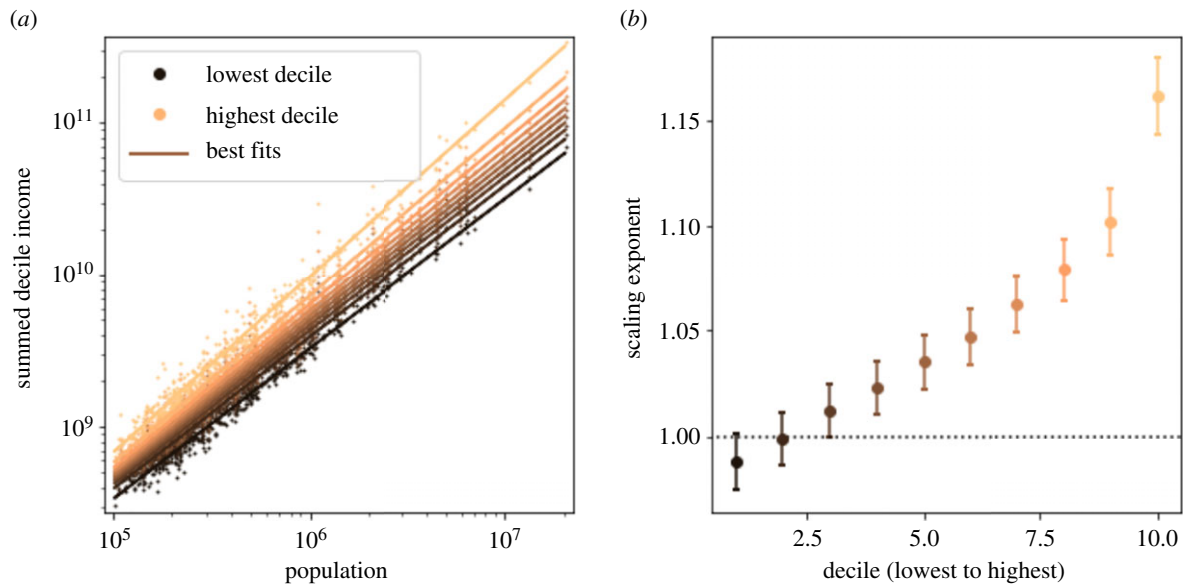Another useful perspective on the scaling of the income distributions is to compare large and small cities using

Figure 3. Scaling of income (in US Dollars) by population for deciles of US cities. (*a*) Scaling of total income in deciles. (*b*) Scaling exponents (*β*) of each decile and corresponding 95% confidence intervals. The dashed line is $\beta = 1$ to help guide the eye. Higher-income deciles exhibit greater scaling exponents than lower income deciles, and the lowest deciles exhibit near-linear scaling. The scaling exponents for aggregated income in city, combining all deciles, is 1.07.
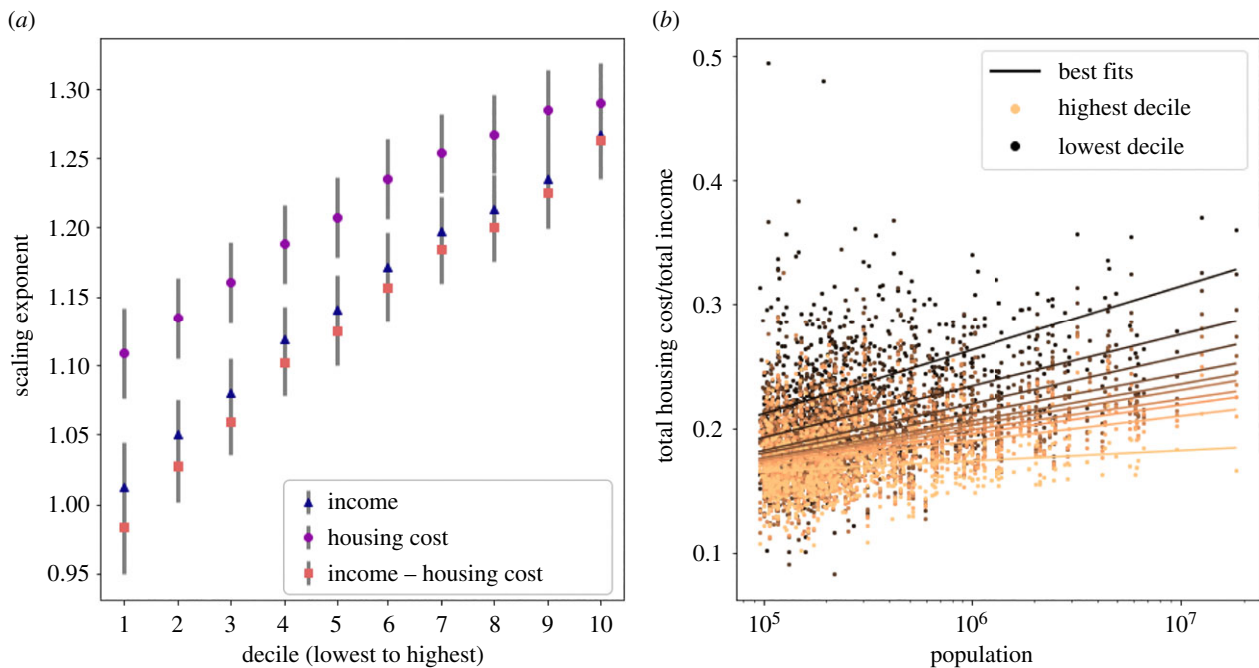

Figure 4. Comparing the scaling of housing cost and income. (*a*) The scaling of total income, total housing cost and the difference between total income and housing cost, for each decile. Housing cost scales with greater exponents than income for all deciles. The housing-adjusted income exhibits similar variation across deciles as total income. (*b*) Ratio between housing cost and household income as a function of city population. In the poorest deciles (dark brown), the proportion of income spent on housing increases sharply with city size; in the wealthiest deciles (orange), this proportion remains stagnant.

measures that consider the entire distribution through the KL divergence. Figure 6 shows the KL divergence between each US city and the largest city, as a function of the log-transformed city population. The KL divergence, in general, decreases with increasing city population, and approaches zero as the population approaches that of the largest city. This behaviour suggests that as cities get smaller, their income distributions are increasingly dissimilar to that of the largest city. The Pearson correlation between the two variables in figure 6 is −0.259, while the Spearman correlation is −0.718. The Pearson correlation measures the linear correlation between two variables, while the Spearman

correlation measures the rank correlation, and assesses how well the relationship between two variables can be described by a monotonic function, regardless of linearity [28]. This finding suggests that population and the KL divergence tend to change together, but not necessarily at a constant rate. While we can identify a general scaling trend, our data also exhibit frequent outliers and deviations.

4. Discussion

Here, we proposed a new method to study the scaling of income distributions and income inequality in urban areas.
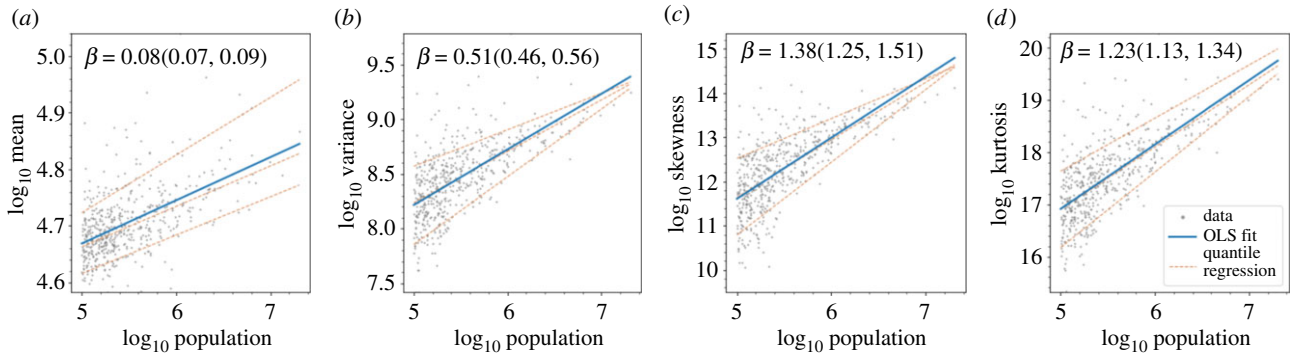
Figure 5. First four statistical moments of the estimated income distributions as a function of city size. The texts in each panel display the scaling exponent, β, obtained from ordinary least squares and in the bracket, corresponding 95% confidence intervals. The dashed lines represent quantile regression results from 10%, 50% and 90% percentiles (see electronic supplementary material, table S3 for quantile estimates of β).
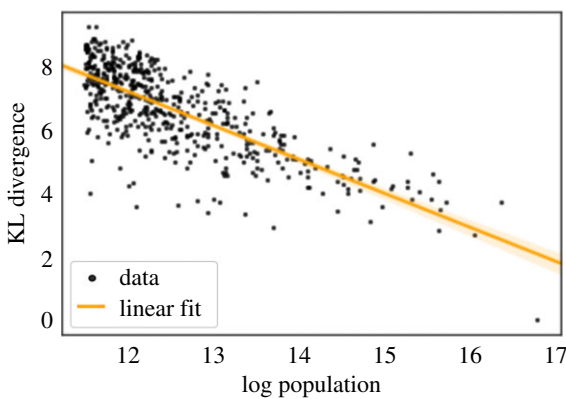

Figure 6. Kullback–Leibler divergence between the estimated income distributions and that of the largest city, as a function of log population. The Spearman correlation is −0.718.

Figure 7. Changes in the Gini coefficient with urban population in simulated log-normal distributions. For a scenario of parallel decile scaling (figure 1*b*) and for a scenario where the deciles have divergent scaling (figure 1*d*). As expected, the divergent scaling observation corresponds to increasing Gini coefficient with population.

The aggregated income in income deciles scales systematically with city size. The bottom decile scales with an exponent slightly below 1 and the top decile with an exponent of $\beta = 1.15$. This result suggests that the benefits of larger cities are increasingly unequally distributed, and for the poorest income deciles, cities have no positive effect over the null expectation of a linear increase. Much has been written about the apparent increasing gains of large cities [4,5], such as greater GDP, higher wages and more patents *per capita*. Our results show that the increasing benefits of city size are not evenly distributed to people within those cities. We further show systematic variations in distribution characteristics. Besides greater mean, distributions of bigger cities also exhibit greater spread, greater asymmetry and heavier tails. These perspectives can be explicitly connected to traditional measures of income inequality, such as the Gini coefficient. Like the Gini coefficient, our method characterizes the overall dispersion of income distributions (figure 7), but it also provides more detailed information that is not characterized by Gini, such as how the urban agglomeration effect alters the incomes of relatively poor or rich people differently.

Since income distribution data represent a sub-sample of the population, it is important to consider how sub-sampling may affect the scaling exponents [29]. We performed a robustness check and found that our conclusions are not affected by the sampling effect described in [29]—see electronic supplementary material for more details.

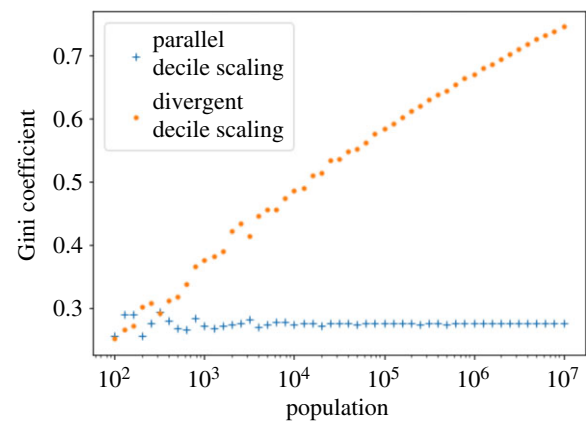Our qualitative conclusions, at first glance, appear to closely align with those of Sarkar *et al.* [26,27], which analyse Australian income data. While both studies observe inequality, the two studies define inequality differently. Sarkar *et al.* uses a non-scale-adjusted method and study income aggregated in income brackets, while we use a scale-adjusted method and aggregate by income deciles. The baseline 'equal' situation is different in the two methods. In Sarkar *et al.*, equality requires the proportion of individuals in each income bracket to remain the same as city size changes, and inequality can appear when the average income increases with city population, even without changes to the shape of the distribution. In our method, equality requires the dispersion in the income distribution to remain the same, regardless of changes in average income. Thus, our methodology allows us to distinguish between increasing income variance and increasing mean income, while Sarkar's does not. Furthermore, the scale-adjusted approach is important when comparing income with other urban indicators, such as housing costs. Since the average of both variables scale with city population, and sometimes with different exponents, it is important to use the scale-adjusted approach to derive meaningful comparisons.

Our paper offers new contributions to the literature. First, we develop a new method to study income inequality in the urban scaling framework, which untangles the systematic shift in mean from the study of income inequality. This method enables us to study how income agglomeration

effects vary between relatively rich and poor people, after accounting for the systematically increasing mean with population size. Second, our analysis including housing cost demonstrates that despite agglomeration effects on income, bigger cities are less affordable for people of all deciles in the sense that they spend proportionally more of their income on housing; this is especially true for lower-income people. Third, our analysis extends beyond the single-parameter characterization of income inequality. We analyse more complex properties of income distributions through analysing statistical moments and KL divergence, and reveal systematic variations with city size. Fourth, our results suggest new directions for understanding mechanisms of urban agglomeration effects—it is important to extend beyond theories considering homogeneous densifying interactions to those which account for heterogeneity.

Understanding the underlying mechanisms of why inequality is systematically scaling with city size is of great future interest with many potential implications. Urban scaling theory in general proposes densifying interactions within cities as the fundamental process leading to the superlinear increase of many features [3,9,10,30,31]. Our analysis shows that the superlinear scaling is not seen within all subsections of the city. The superlinear scaling of total wealth is driven by the top income deciles, and is not matched proportionally by the lowest deciles. This adds another dimension to considerations of the underlying mechanisms of urban scaling theory: what processes are leading to the increasingly unequal distribution of wealth in larger cities? We explored the idea of city heterogeneity as an indirect proxy for heterogeneous interaction rates. One hypothesis of the mechanism driving superlinear scaling of income with city size is that larger cities foster more and more diverse social and economic interactions, creating opportunities for the exchange of ideas and resources. Existing literature credits superlinear growth of income in cities to more opportunities for social contacts and interactions in large cities [4,9]. Increased social contact with city size has been empirically confirmed [32], and ties between individual's exposure to diverse social connections and economic outcomes have been shown empirically as well [33]. Together, this seems to suggest that cities that are better mixed either physically or virtually, allowing diverse parts of the population to be exposed to one another, should be overperforming with respect to urban scaling. We hypothesize that cities with high levels of economic segregation, inhibiting mixing between diverse populations, will underperform with respect to income scaling. Our finding encourages future work to consider heterogeneous models of interactions, as those clustered in space or social/work circles, to form a more coherent understanding of urban scaling.

References

1. United Nations. 2018 *2018 Revision of world urbanization prospects*. New York, NY: United Nations.

2. Lobo J *et al.* 2020 Urban science: integrated theory from the first cities to sustainable metropolises. *Report submitted to the NSF on the Present State and Future of Urban Science*. See https://ssrn.com/abstract=3526940.

3. Pumain D, Paulus F, Vacchiani-Marcuzzo C, Lobo J. 2006 An evolutionary theory for interpreting urban scaling laws. *Cybergeo: Eur. J. Geogr.* document 343. (doi:10.4000/cybergeo.2519)

4. Bettencourt LM, Lobo J, Helbing D, Kühnert C, West GB. 2007 Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306. (doi:10.1073/pnas.0610172104)

5. Bettencourt LM, Lobo J. 2016 Urban scaling in Europe. *J. R. Soc. Interface* **13**, 20160005. (doi:10.1098/rsif.2016.0005)

6. Meirelles J, Neto CR, Ferreira FF, Ribeiro FL, Binder CR. 2018 Evolution of urban scaling: evidence from Brazil. *PLoS ONE* **13**, e0204574. (doi:10.1371/journal.pone.0204574)

7. Zünd D, Bettencourt LMA. 2019 Growth and development in prefecture-level cities in China. *PLoS ONE* **14**, e0221017. (doi:10.1371/journal.pone.0221017)

8. Sahasranaman A, Bettencourt LMA. 2019 Urban geography and scaling of contemporary Indian cities. *J. R. Soc. Interface* **16**, 20180758. (doi:10.1098/rsif.2018.0758)

9. Bettencourt LM. 2013 The origins of scaling in cities. *Science* **340**, 1438–1441. (doi:10.1126/science.1235823)

10. Yang VC, Papachristos AV, Abrams DM. 2019 Modeling the origin of urban-output scaling laws. *Phys. Rev. E* **100**, 032306. (doi:10.1103/PhysRevE.100.032306)

11. Bowles S, Fong CM, Gintis H, Pagano U. 2012 *The new economics of inequality and redistribution*. Cambridge, MA: Cambridge University Press.

12. Piketty T. 2014 *Capital in the 21st Century*. Cambridge, MA: Harvard University Press.

13. Persson T, Tabellini G. 1991 *Is inequality harmful for growth?: Theory and evindence*. Cambridge, MA: Queen Elizabeth House.

14. Alesina A, Rodrik D. 1994 Distributive politics and economic growth. *Q. J. Econ.* **109**, 465–490. (doi:10.2307/2118470)

15. Li H, Zou H-F. 1998 Income inequality is not harmful for growth: theory and evidence. *Rev. Dev. Econ.* **2**, 318–334. (doi:10.1111/1467-9361.00045)

16. Forbes KJ. 2000 A reassessment of the relationship between inequality and growth. *Am. Econ. Rev.* **90**, 869–887. (doi:10.1257/aer.90.4.869)

17. Barro RJ. 2000 Inequality and growth in a panel of countries. *J. Econ. Growth* **5**, 5–32. (doi:10.1023/A:1009850119329)

18. Pagano P. 2004 An empirical investigation of the relationship between inequality and growth. Economic working papers 536, Bank of Italy, Economic Research and International Relations Area. See https://ideas.repec.org/p/bdi/wptemi/td_536_04.html.

19. Alesina A, Perotti R. 1996 Income distribution, political instability, and investment. *Eur. Econ. Rev.*

40, 1203–1228. (doi:10.1016/0014-2921(95) 00030-5)

20. Kelly M. 2000 Inequality and crime. *Rev. Econ. Stat.* **82**, 530–539. (doi:10.1162/003465300559028)

21. Jong-Sung Y, Khagram S. 2005 A comparative study of inequality and corruption. *Am. Sociol. Rev.* **70**, 136–157. (doi:10.1177/000312240507000107)

22. Glaeser EL, Resseger M, Tobio K. 2009 Inequality in cities. *J. Reg. Sci.* **49**, 617–646. (doi:10.1111/j.1467-9787.2009.00627.x)

23. Behrens K, Nicoud FR. 2014 Survival of the fittest in cities: urbanisation and inequality. *Econ. J.* **124**, 1371–1400. (doi:10.1111/ecoj.12099)

24. Cottineau C, Finance O, Hatna E, Arcaute E, Batty M. 2019 Defining urban clusters to detect agglomeration economies. *Environ. Plann. B: Urban Anal. City Sci.* **46**, 1611–1626. (doi:10.1177/2399808318755146)

25. Snow NB, Pavan R. 2013 Inequality and city size. *Rev. Econ. Stat.* **95**, 1535–1548. (doi:10.1162/REST_a_00328)

26. Sarkar S, Phibbs P, Simpson R, Wasnik S. 2018 The scaling of income distribution in Australia: possible relationships between urban allometry, city size, and economic inequality. *Environ. Plann. B: Urban Anal. City Sci.* **45**, 603–622. (doi:10.1177/0265813516676488)

27. Sarkar S. 2019 Urban scaling and the geographic concentration of inequalities by city size. *Environ. Plann. B: Urban Anal. City Sci.* **46**, 1627–1644. (doi:10.1177/2399808318766070)

28. Salkind N. 2007 *Encyclopedia of measurement and statistics*. Thousand Oaks, CA: SAGE.

29. Gómez-Liévano A, Vysotsky V, Lobo J. 2021 Artificial increasing returns to scale and the problem of sampling from lognormals. *Environ. Plann. B: Urban Anal. City Sci.* **48**, 1574–1590. (doi:10.1177/2399808320942366)

30. Gomez-Lievano A, Patterson-Lomba O, Hausmann R. 2016 Explaining the prevalence, scaling and variance of urban phenomena. *Nat. Human Behav.* **1**, 1–6. (doi:10.1038/s41562-016-0012)

31. Pan W, Ghoshal G, Krumme C, Cebrian M, Pentland A. 2013 Urban characteristics attributable to density-driven tie formation. *Nat. Commun.* **4**, 1–7. (doi:10.1038/ncomms2961)

32. Schläpfer M, Bettencourt LM, Grauwin S, Raschke M, Claxton R, Smoreda Z, West GB, Ratti C. 2014 The scaling of human interactions with city size. *J. R. Soc. Interface* **11**, 20130789. (doi:10.1098/rsif.2013.0789)

33. Chetty R, Hendren N. 2018 The impacts of neighborhoods on intergenerational mobility I: childhood exposure effects. *Q. J. Econ.* **133**, 1107–1162. (doi:10.1093/qje/qjy007)