scientific reports

OPEN

Commuting network effect on urban wealth scaling

Luiz G. A. Alves [1], Diego Rybski [2,3,4] & Haroldo V. Ribeiro [5]

Urban scaling theory explains the increasing returns to scale of urban wealth indicators by the per capita increase of human interactions within cities. This explanation implicitly assumes urban areas as isolated entities and ignores their interactions. Here we investigate the effects of commuting networks on the gross domestic product (GDP) of urban areas in the US and Brazil. We describe the urban GDP as the output of a production process where population, incoming commuters, and interactions between these quantities are the input variables. This approach significantly refines the description of urban GDP and shows that incoming commuters contribute to wealth creation in urban areas. Our research indicates that changes in urban GDP related to proportionate changes in population and incoming commuters depend on the initial values of these quantities, such that increasing returns to scale are only possible when the product between population and incoming commuters exceeds a well-defined threshold.

Cities are engines of innovation[1], happiness[2], creativity[3], production of knowledge[4] and wealth[5], but are also associated with the spread of diseases[6], violent crimes[7–9], segregation[10,11], and pollution[12,13]. The prominent role of urban systems in the society and the increasing number of people living in cities have driven significant and ongoing efforts towards developing a new science of cities[12,14–16], which has been further fostered by the increasing availability of highly detailed city data.

Urban scaling is among the most remarkable and widely acclaimed findings of this new science of cities[3]. This theory states that different urban indicators (Y) can be expressed by a power-law function of city population (N), that is, $Y \sim N^{\beta}$, where β is the urban scaling exponent or the scale-invariant elasticity. A direct consequence of this nonlinear relationship is that percentage increments in population may result in greater or smaller percentage changes in urban indicators depending on whether $\beta > 1$ (increasing returns to scale) or $\beta < 1$ (decreasing returns to scale). The latter case is well illustrated by economies of scale in material infrastructure of large cities[3], while the more than proportional increase of urban wealth with city population is a typical example of the former case[3,17].

In an analogy with the origins of allometric scaling in biology[18], the emergence of these urban scaling laws has been attributed to intra-city networks[19]. The underlying infrastructure network of cities (such as electric cable and road networks) plays the role of the biological networks (such as circulatory and respiratory networks), and network constraints would explain the decreasing returns to scale observed in infrastructure indicators as well as in animals metabolic rates. However, this analogy is limited as there is no biological counterpart for the emergence of increasing returns to scale in urban socioeconomic indicators. The per capita increase of social and economic quantities with city population is usually attributed to the superlinear scaling of human interaction networks within cities[19–22], that is, to the finding that per capita levels of individual contacts and communication activities also increase with population.

Urban scaling theory implicitly assumes cities as isolated entities and that only within-city interactions would be responsible for the increasing returns to scale of city socioeconomic indicators. This assumption contrasts with the complex inter-city interactions[23–31] that manifest through people's everyday movements from home to work (commuting networks) and long-term migration patterns among cities. These networks allow cities to attract people from different neighboring areas, contribute to information flow, city development, and the formation of more creative and productive teams[32]. Despite the importance of inter-city interactions, there have been very few attempts to put forward the more holistic idea that urban scaling may not only emerge from processes occurring within the city boundaries[33].

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA. [2]Potsdam Institute for Climate Impact Research – PIK, Member of Leibniz Association, P.O. Box 601203, 14412 Potsdam, Germany. [3]Department of Environmental Science Policy and Management, University of California Berkeley, 130 Mulford Hall #3114, Berkeley 94720, CA, 94720, USA. [4]Complexity Science Hub Vienna, Josefstädterstrasse 39, A-1090, Vienna, Austria. [5]Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil. ✉email: ca-dr@rybski.de; hvr@dfi.uem.br
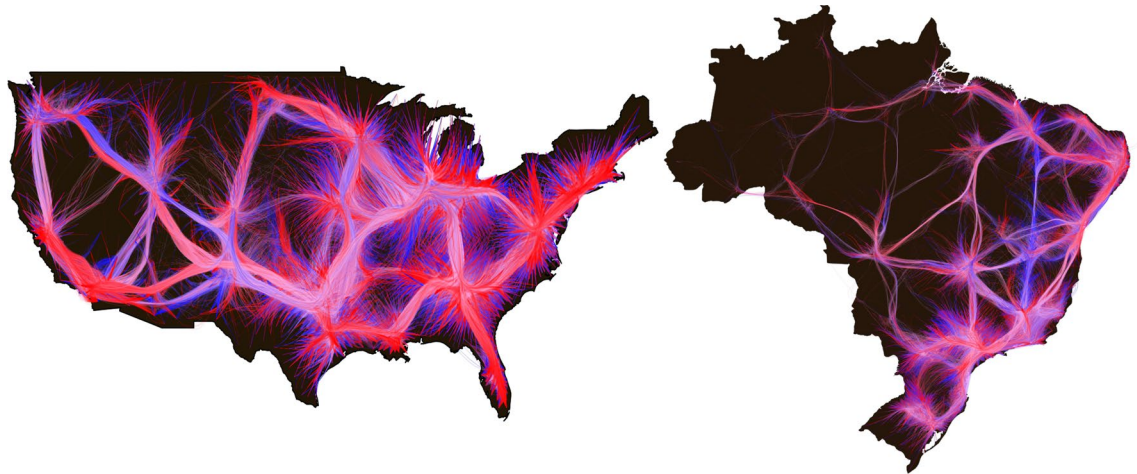
Figure 1. Visualization of the commuting networks among US counties (left) and Brazilian (right) municipalities. In this representation, nodes are urban areas and links represent the flow of commuters among them. Edge widths are associated with the flow intensity and colors refer to flow direction (bluish tones indicate outgoing links while reddish indicate incoming connections). This figure has been made using matplotlib[35] and geopandas[36], we have further aggregated links based on their spatial proximity using a kernel-based edge bundling algorithm[37]. The overall complexity of interactions among urban areas motivates investigations about the role of these networks on urban scaling.

Here we investigate the effects of commuting networks on the urban gross domestic product (GDP) scaling using data from the US and Brazil. To do so, we use an analogy with the economic theory of production functions[34] proposed by Ribeiro et al.[13] to account for the interconnected role of population and incoming commuters on urban wealth creation. Our results show that modeling urban GDP as a function of population and number of incoming commuters significantly improves the description of urban productivity when compared with usual urban scaling. These models explicitly consider interactions between population (N) and number of incoming commuters (S) and indicate that the elasticity of scale of urban GDP depends on the product of these two quantities ($\Omega = N \times S$). This dependence implies the existence of a transition from decreasing to increasing returns to scale when Ω exceeds a particular threshold. Our research suggests that cities increase their productivity from a combination of intra- and inter-city processes but that increasing returns to scale of urban GDP only emerge when the intensity of the product of these processes exceeds a threshold. Therefore, the GDP of two cities with similar populations may respond differently to a percentage change in population if their number of incoming commuters is different. The comparison between the results for the US and Brazil further suggests that the elasticity of scale also depends on the development degree of the urban area.

Results

We start by investigating the complex network related to people's everyday movements from home to work. We have collected data from the US Census Bureau and the Brazilian Institute of Geography and Statistics about the commuting flow among urban areas in the US and Brazil (see "Methods" for details). The urban units comprise 3,220 US counties and 5,565 Brazilian municipalities. Figure 1 shows these spatial networks where vertices are urban areas (US counties or Brazilian municipalities), and links represent the flow of workers among these urban areas. Links are directed from home to workplace, and their weights stand for the total number of workers along a given path. The complexity of these commuting networks contrasts with the current view of urban scaling theory that assumes cities as isolated entities and intra-city processes as the only ones responsible for urban scaling. It seems natural to assume that the flow of people among urban areas contributes, at least to a certain extent, to wealth and knowledge creation in cities; however, this possibility remains little explored[33].

Before we approach this problem, let us first revisit the urban scaling of GDP. To do so, we have also obtained population and GDP data for all US counties and Brazilian municipalities in our data set (see "Methods" for details). The urban scaling (also known as the Bettencourt-West law)[3,15] states that GDP (Y) and population (N) of urban areas are related via a power-law function

$$Y \sim N^{\beta} \quad \text{or} \quad \log Y \sim \beta \log N, \tag{1}$$

where β is the urban scaling exponent. Furthermore, urban scaling theory predicts β to be larger than one for GDP and other socioeconomic indicators[19]. This would thus explain the agglomeration economies in cities, that is, the more than proportional increase of urban wealth with city population. We estimate the values of $\beta = 1.04 \pm 0.003$ for US counties and $\beta = 1.04 \pm 0.01$ for Brazilian municipalities via usual least-squares method (see Fig. S1). These values are statistically larger than one ($p-$values $= 0$, permutation test) and agree with urban scaling theory in the sense that socioeconomic indicators are expected to display a superlinear relation with city population[3,19]. These results further indicate that a 1% increase in the population of an urban area implies an
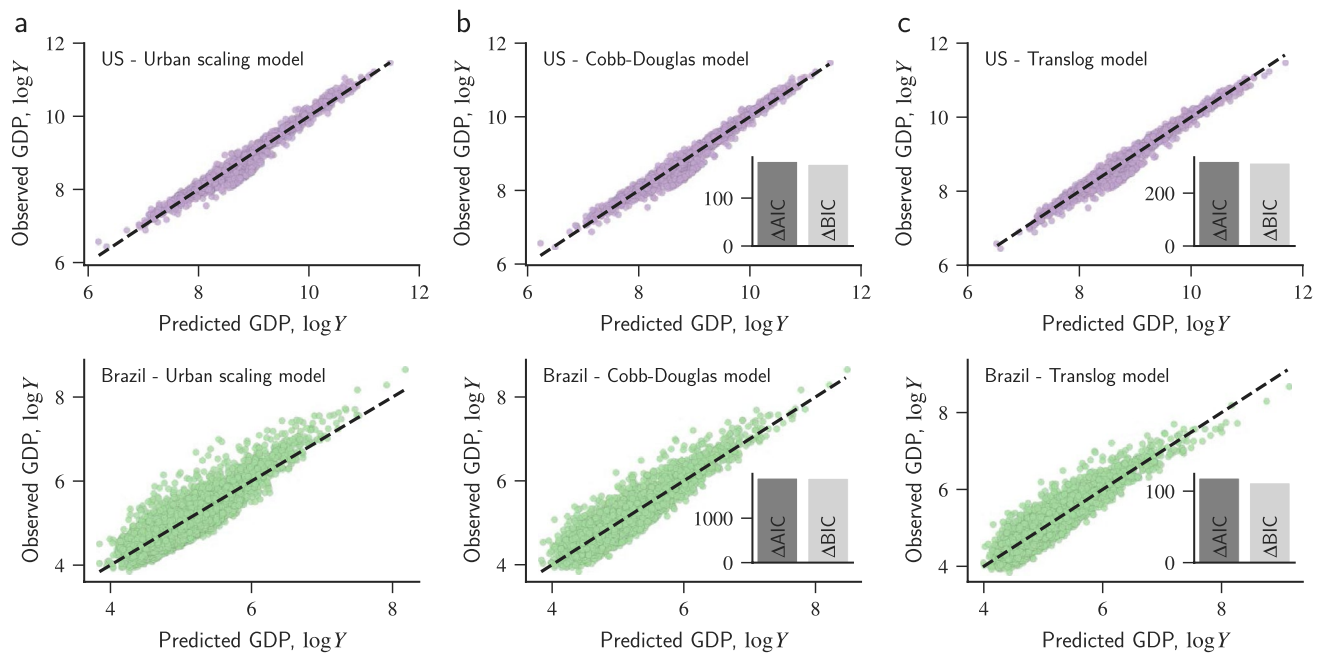
Figure 2. Urban scaling predictions of GDP and the improved descriptions based on the interplay between population and incoming commuters. Relation between the observed values of urban GDP and those predicted by (**a**) usual urban scaling (Eq. 1), (**b**) Cobb-Douglas model (Eq. 2), and (**c**) translog model (Eq. 3). All quantities are expressed in base-10 logarithmic scale and the dashed lines represent the 1:1 relationship. The upper plots (purple markers) show the results for US counties, and the lower ones (green markers) represent the values for Brazilian municipalities. The bar plot inserted in panel (b) shows the differences in the Akaike information criterion (△AIC) and Bayesian information criterion (△BIC) estimated between the Cobb-Douglas and urban scaling models. The large positive values of these quantities indicate that the Cobb-Douglas model significantly improves the description of urban GDP when compared with usual urban scaling. The inset in panel (**c**) shows the same quantities (△AIC and △BIC) calculated between the translog and the Cobb-Douglas models, where the large positive values indicate that the translog model further refines the goodness of fit.

increase of 1.04% in its GDP, both for the US and Brazil. Figure 2a shows the relationship between the observed values of urban GDP and those predicted by urban scaling (Eq. 1).

Despite the good quality of the predictions based on Eq. (1), we observe some systematic deviations as urban scaling tends to underestimate the urban GDP for large values of this quantity (especially for Brazilian municipalities). Furthermore, the universality of these scaling exponents has been debated, and results indicate that their values vary from country to country[33], are affected by spatial distributions of cities[31], and susceptible to different city definitions[38–40]. In this regard, it is worth noticing that the exponent reported here for US counties is smaller than the results obtained when considering the urban units defined in terms of US metropolitan statistical areas[19] and close to the value obtained when defining US cities as functional urban areas[33]. The combination of these caveats with the overall complexity of commuting networks thus invites more complex modeling approaches of urban GDP that account for additional dependent variables beyond urban population. In this context, an interesting analogy with the economic theory of production functions[34] has been recently proposed for modeling carbon dioxide emissions while simultaneously considering the effects of population and area[13].

Inspired by the approach of Ref.[13], we propose to model urban GDP (Y) as the output of a two-factor production model where population (N) and incoming commuters (S) in a given urban area are the inputs of this hypothetical production process. Mathematically, we write $Y = F(N, S)$, where F represents the form of the production function that we borrow from the theory of production functions[34]. Just like combinations of capital and labor may lead to the similar output of goods in a factory, our approach thus considers that a city generates wealth through a process that is mediated by its population (intra-city process) and by the population of workers that is attracted from other cities (inter-city process). It is worth noticing that the number of incoming or outgoing commuters are the simplest features we can extract from commuting networks (in network language they represent the in- and out-strength of nodes[41], respectively). Beyond a simplifying condition, our choice for considering only the number of incoming commuters as an input factor of the production function relates to the fact that this variable can be more directly associated with cities' importance and development degree. We expect a city that attracts many workers from neighboring cities to have a more diverse and developed labor market than a counterpart with few incoming commuters. A city with a large number of incoming commuters is also likely to affect and exert a large influence on its neighborhood, such that this variable can be thought of as an indirect proxy for quantifying the degree of interaction among cities. We further believe the incoming commuters variable fits nicely with the analogy of production functions and allows a direct interpretation of substitution effects between internal and commuting populations on urban wealth production. Moreover, we

have verified that the numbers of incoming and outgoing commuters are strongly correlated, and the inclusion of outgoing commuters in the models we shall discuss does not statistically improve the description of urban GDP.

We first consider the Cobb-Douglas model[42]

$$Y \sim N^{\beta_N} S^{\beta_S} \quad \text{or} \quad \log Y \sim \beta_N \log N + \beta_S \log S, \qquad (2)$$

where β_N and β_S are two independent parameters. This model is one of the most widely used production functions[34] and can be thought of as a generalization of Eq. (1) that is recovered by setting $\beta_N = \beta$ and $\beta_S = 0$. Similarly to usual urban scaling, the Cobb-Douglas production function has a scale-invariant elasticity of scale equal to $\varepsilon = \beta_N + \beta_S$[34]. This means a proportionate change in population (N) and incoming commuters (S) associates with a proportionate change in urban GDP independent of N and S. Moreover, we have increasing returns to scale when $\beta_N + \beta_S > 1$ (doubling N and S implies more than doubling Y), decreasing returns to scale when $\beta_N + \beta_S < 1$ (doubling N and S implies less than doubling Y), and constant returns to scale when $\beta_N + \beta_S = 1$ (doubling N and S implies doubling Y).

We can interpret Eq. (2) as the result of including incoming commuters as a confounding variable in the usual urban scaling (Eq. 1) via a multiple linear regression in the log-transformed variables. Thus, if the Cobb-Douglas model represents a significantly better description of data, we cannot reject the hypothesis that incoming commuters contribute to the deviations observed in the usual urban scaling and ultimately to the GDP of an urban area. We have fitted the linearized version of Eq. (2) to data via the ridge regression approach[43,44]. As detailed in "Methods", using this regularization method is essential for accounting for the multicollinearity between the predictor variables ($\log N$ and $\log S$) and yielding stable estimates for the parameters β_N and β_S. This approach yields $\beta_N = 0.93 \pm 0.01$ and $\beta_S = 0.10 \pm 0.01$ for US counties, and $\beta_N = 0.75 \pm 0.01$ and $\beta_S = 0.33 \pm 0.01$ for Brazilian municipalities. All parameters are statistically different from zero in both data sets ($p-$values < 0.05, permutation test). Figure 2b compares the observed values of urban GDP for US counties and Brazilian municipalities with the predictions of the Cobb-Douglas model. We note that this model improves the description of urban GDP for both countries and reduces the bias for large GDP values observed for usual urban scaling. Furthermore, the Akaike and Bayesian information criteria (insets of Fig. 2b) indicate that it is much more likely that the empirical data come from the Cobb-Douglas model (Eq. 2) than from the usual urban scaling relation (Eq. 1).

The significantly better description of data by the Cobb-Douglas model and the fact that $\beta_N < \beta$ for both data sets indicates that urban scaling is considerably affected by the confounding factor of incoming commuters. Moreover, we find $\beta_N + \beta_S > 1$ for both countries such that every 1% increase in population and incoming commuters associates with 1.03% and 1.08% raises in the GDP of US counties and Brazilian municipalities, respectively. While the precise comparison between the results of both countries demands a unified city definition (which is not our case), the significantly different values of β_S suggest that the GDP of urban areas in Brazil is more affected by commuters than in the US case. Indeed, the Cobb-Douglas model predicts that a 1% increase in incoming commuters (keeping population unaltered) implies a 0.33% rise in the GDP of Brazilian municipalities, while the GDP of US counties is expected to increase only 0.10%. This result suggests that the US urban labor market is significantly more integrated than the Brazilian one; that is, the US commuting network is in a much more optimized state such that changes in the number of commuters have a small effect on urban GDP. On the other hand, the developing stage of Brazil is likely to impose considerable infrastructure limitations that hamper its commuting network to achieve a similar optimal state. These limitations are likely to disproportionately affect small urban areas rather than large urban centers, which in turn may explain the more significant dependence of urban GDP on incoming commuters observed in Brazil. Thus, the larger value of β_S for Brazil may represent a latent possibility for increasing the overall GDP of Brazilian municipalities (especially the small ones) that seems limited by transportation infrastructure.

Although the Cobb-Douglas model represents a better description of our data, it also imposes some limitations to the productivity of cities (as well as to firms[34]) that may not necessarily hold. One of these limitations is associated with the elasticity of scale (ε) independent of population and incoming commuters. Another one is related to the so-called elasticity of substitution[34]. This quantity represents the ratio between a proportionate change in the inputs and the associated proportional change in the slope of an isoquant of $F(N, S)$, that is, the elasticity of substitution measures the efficiency at which we can substitute population by incoming commuters while keeping the same level of GDP. The elasticity of substitution is independent on N, S, and Y, and always unitary in the Cobb-Douglas model[34].

To overcome these limitations and introduce more flexibility to our modeling of urban GDP, we have considered the transcendental logarithmic (translog) production function[34,45]

$$\log Y \sim \beta_N \log N + \beta_S \log S + \beta_C \log N \log S, \qquad (3)$$

where β_C is an additional parameter. We observe that the translog model generalizes the Cobb-Douglas function (recovered by setting $\beta_C = 0$) by including an interaction term between population and incoming commuters. This interaction allows population effects on GDP to depend on the number of commuters as well as commuters effects on GDP to depend on population. Indeed, a small proportionate increment in population [$N \to (1 + x)N$, with $x \ll 1$] implies a proportionate increment in GDP that depends on incoming commuters [$Y \to (1 + x(\beta_N + \beta_C \log S))Y$]. Similarly, a small proportionate increment in incoming commuters [$S \to (1 + x)S$] also results in a proportionate increment in GDP that depends on population [$Y \to (1 + x(\beta_S + \beta_C \log N))Y$]. The translog model has an elasticity of scale $\varepsilon = \beta_N + \beta_S + \beta_C \log(NS)$[34] that depends on population and incoming commuters. Thus, variations in GDP associated with proportionate changes in population and incoming commuters depend not only on the magnitude of these changes but also on the initial values of N and S. All these features are not present in usual urban scaling (Eq. 1) nor in Cobb-Douglas model

(Eq. 2), as both models implicitly assume changes in urban GDP to be scale invariant. The translog model relaxes this hypothesis and adds more flexibility to describe urban GDP in terms of population and incoming commuters.

We have fitted the translog model to our data with the ridge regression approach (see "Methods" for details), finding $\beta_N = 0.76 \pm 0.01$, $\beta_S = -0.14 \pm 0.01$, and $\beta_C = 0.05 \pm 0.002$ for US counties, and $\beta_N = 0.62 \pm 0.01$, $\beta_S = -0.01 \pm 0.03$, and $\beta_C = 0.08 \pm 0.01$ for the Brazilian municipalities. Except for β_S in the Brazilian data set, all model coefficients are statistically different from zero (p−values < 0.05, permutation test). It is worth noting that $\beta_S \approx 0$ does not mean incoming commuters have no role on the GDP of Brazilian municipalities, but that the interaction term related to population and incoming commuters is more significant than the isolated effect of commuters. On the other hand, the negative value of β_S obtained for the US data indicates that increasing the number of incoming commuters has a negative effect on the GDP of small counties. Indeed, by rewriting Eq. (3) as $\log Y \sim \beta_N \log N + (\beta_S + \beta_C \log N) \log S$, one readily notes that an increase in S associates with a reduction in Y when $N < N^*$ with $N^* = 10^{-\beta_S/\beta_C}$. We have $N^* \approx 631$ for US counties, which in turn allows a negative association between GDP and incoming commuters for tiny urban areas. In practice, this negative association would be limited to the eleven US counties having population smaller than 631 people.

Figure 2c shows the relation between the actual values of urban GDP for the US and Brazil and the predictions based on the translog model. We have verified that the translog model (Eq. 3) improves the description of urban GDP in both data sets when compared with the usual urban scaling (Eq. 1) and the Cobb-Douglas production function (Eq. 2), reducing even further the biases observed for usual urban scaling. The Akaike and Bayesian information criteria (insets of Fig. 2c) indicate it is much more likely that the empirical data come from the translog model than from the Cobb-Douglas model; consequently, data is even more likely to come from the translog model than from usual urban scaling.

In addition to representing a significantly better description for the GDP of US counties and Brazilian municipalities, the adjusted translog models indicate that the effects of population and incoming commuters on urban GDP intensifies with the increase of the product $\Omega = NS$ (that is, $\beta_C > 0$). The elasticity of scale is $\varepsilon = 0.62 + 0.05 \log(NS)$ for US counties and $\varepsilon = 0.61 + 0.08 \log(NS)$ for Brazilian municipalities. Thus, the greater the population and incoming commuters of a city, the larger the impact of proportionate changes of these quantities on its urban GDP. This effect is also more intense for Brazilian municipalities than US counties, such that urban areas of both countries with similar populations and incoming commuters respond differently to proportionate changes in these quantities, with urban areas in Brazil reporting proportionally more GDP gain than the US ones. It is worth noticing that this behavior is not related to raw values of urban GDP (which are almost always higher in the US than in Brazilian urban areas); instead, elasticity of scaling is associated with proportionate changes in GDP. A similar result was observed in the context of usual urban scaling of GDP for West and East cities of the European Union, where West cities (usually richer and more developed than their East counterparts) display smaller increasing returns to scale than East cities[46]. Again, we can associate this behavior with the developing nature of Brazilian municipalities in contrast with the more steady economic situation of US counties.

Intriguingly, the translog model predicts that urban areas transit from decreasing to increasing returns to scale of GDP depending on the values of N and S. By analyzing the inequality $\varepsilon > 1$, we find that whether the product between population and incoming commuters $\Omega = NS$ exceeds or not the critical value $\Omega^* = 10^{(1-\beta_N-\beta_S)/\beta_C}$ determines if an urban area shows decreasing ($\Omega < \Omega^*$) or increasing ($\Omega > \Omega^*$) returns to scale[13]. Moreover, the line defined by $\log S = \log \Omega^* + \log N$ separates two regions in a plane of $\log S$ versus $\log N$, such that urban areas with increasing returns to scale are above this line, and those with decreasing returns to scale are below this line. Figure 3a shows this plane in which the color-coded background refers to the elasticity of scale for each point in the plane and the gray circles stand for values of $\log S$ versus $\log N$ for all US counties and Brazilian municipalities.

We note that incoming commuters and population are approximately power-law related, that is, $S \sim N^\gamma$ (or $\log S \sim \gamma \log N$) with the exponent $\gamma = 0.87 \pm 0.01$ for the US and $\gamma = 1.06 \pm 0.01$ for Brazil. However, the spread around this power-law trend allows the elasticity of scale to effectively depend on both population and incoming commuters. Thus, urban areas with similar populations (as illustrated for the US urban areas Bethel-AK and Montour-PA, and the Brazilian municipalities Marzagão-GO and Monções-SP) can have significantly different elasticity of scale. Analogously, urban areas with similar number of incoming commuters (as illustrated for the US urban areas McMullen-AK and Hawaii-HI, and the Brazilian municipalities Serra da Saudade-MG and Sant'Ana do Livramento-RS) can also display significantly different elasticity of scale. It is also worth mentioning that the value of $\gamma > 1$ observed for Brazil agrees with the idea that the developing stage of the Brazilian transportation infrastructure disproportionately affects small urban areas, creating a more uneven distribution of incoming commuters among Brazilian municipalities. In contrast, the $\gamma < 1$ observed for the US indicates a more balanced distribution of incoming commuters and that even small urban areas can have a significant number of incoming commuters.

Using the expression for ε and the adjusted parameters, we calculate the elasticity of scale for every county and municipality in our data set. Figure 3b depicts the probability distribution of this quantity for the US and Brazil. While both countries have a similar number of urban areas with $\varepsilon > 0$ (approximately 63% of US counties and Brazilian municipalities), the probability distribution of ε for Brazilian municipalities is considerably more skewed to the right. Because of this, the average elasticity of scale of Brazilian urban areas ($\langle \varepsilon \rangle = 1.04 \pm 0.001$) is slightly larger than those in the US ($\langle \varepsilon \rangle = 1.02 \pm 0.001$), and the number of Brazilian urban areas with more extreme values of ε is significantly larger than in the US (for instance, Brazil has 262 municipalities with $\varepsilon > 1.2$ while only 34 US counties meet this criterion).

Despite the differences between the definitions of usual urban scaling exponent β and elasticity of scale ϵ, both quantities measure the response of GDP under changes in population (β) and in population and incoming commuters (ϵ), such that a qualitative comparison between these quantities is possible. Having made this consideration, the fact that the average elasticity of scale of Brazil is larger than the US one agrees with the negative
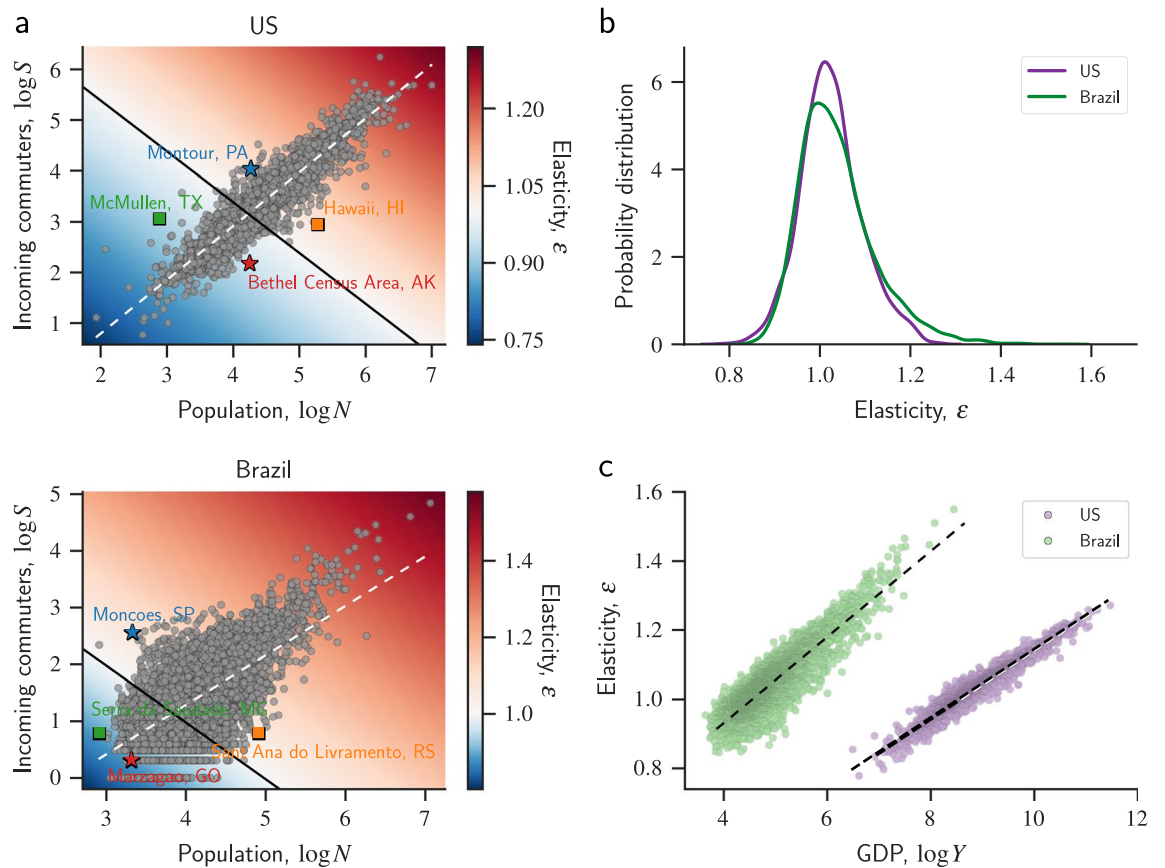
Figure 3. Elasticity of scale and the transition from decreasing to increasing returns to scale of GDP. (**a**) Dependence of the elasticity of scale (ε) on population (N) and incoming commuters (S) calculated from the translog model for US counties (upper panel) and Brazilian municipalities (lower panel). The color-coded background represents the elasticity of scale for each location in the plane population versus incoming commuters (on base-10 logarithmic scale) and the gray circles indicate the values of $\log N$ versus $\log S$ for each US and Brazilian urban area. The solid lines in both panels separate the region of decreasing returns to scale ($\varepsilon < 1$: bluish region) from the increasing returns to scale one ($\varepsilon > 1$: reddish region). We notice that urban areas are spread around the trend defined by $S \sim N^\gamma$ (or $\log S \sim \gamma \log N$, with $\gamma = 0.87 \pm 0.01$ for the US and $\gamma = 1.06 \pm 0.01$ for Brazil – white dashed lines) and that about 63% of urban areas in the US and Brazil display increasing returns to scale. The spread of the relationship between $\log S$ and $\log N$ allows urban areas with similar population sizes (as illustrated by the star markers) or similar incoming commuters (as illustrated by the square markers) to have significantly different values of elasticity of scale. (**b**) Probability distribution of the elasticity of scale estimated for all urban areas in the US (purple) and Brazil (green). The average elasticity of scale $\langle \varepsilon \rangle$ is slightly larger for Brazil than the US ($\langle \varepsilon \rangle = 1.04 \pm 0.001$ and $\langle \varepsilon \rangle = 1.02 \pm 0.001$, respectively), while the Brazilian distribution is significantly more skewed to the right when compared with the US case (skewness $= 1.1$ and skewness $= 0.3$, respectively). (**c**) Relationship between the elasticity of scale and the GDP of urban areas in the US (purple) and Brazil (green). The GDP of Brazilian municipalities has been converted from Brazilian Real (BRL) to their equivalent values in US Dollars (USD) in this plot. We notice that the elasticity of scale is approximately proportional to the logarithm of urban GDP (as indicated by the dashed lines) with the constant of proportionality (κ) larger for Brazil than the US ($\kappa = 0.124 \pm 0.001$ and $\kappa = 0.099 \pm 0.001$, respectively). We also observe that the elasticity of scale of urban areas with similar GDP is always larger for Brazilian municipalities than US counties.

association between the usual urban scaling exponent and GDP of countries reported in Ref.[33] (see Fig. 9 in their supporting information). However, we observe a positive association between elasticity of scale and urban GDP for urban areas in Brazil and US. Figure 3c shows that the elasticity of scale is approximately proportional to the logarithm of the urban GDP ($\varepsilon \sim \kappa \log Y$, with κ being a constant of proportionality) for both countries but with Brazil having a slightly sharper association ($\kappa = 0.124 \pm 0.001$ for Brazil and $\kappa = 0.099 \pm 0.001$ for US). It is worth observing that this association can be explained by the relations $S \sim N^\gamma$ and $Y \sim N^\beta$. By solving these latter relations for N and S and plugging the results into the definition of ε, we find $\varepsilon \sim (\beta_C/\beta)(1 + \gamma) \log Y$ which in turn implies $\kappa = (\beta_C/\beta)(1 + \gamma)$. The estimates of κ using this expression ($\kappa \approx 0.09$ for the US and $\kappa \approx 0.16$ for Brazil) are in reasonably good agreement with the values obtained by fitting the relationship shown Fig. 3c with deviations explained by the estimation errors of the other parameters (β, β_C, and γ) and by the systematic underestimation of S by the adjusted relation with N ($S \sim N^\gamma$) in large Brazilian municipalities.

While appearing contradictory at first glance, the differences in the association between β and the GDP of countries and the association between ε and the GDP of urban areas are likely to represent an aggregation effect. As shown in Fig. 3c, Brazilian municipalities with GDP values similar to US counties always have large values of elasticity of scale, which in turn yields an average elasticity of scale larger for Brazil than US. Thus, we expect the elasticity of scale to be associated with the development degree of the urban area as also discussed by Strano et al.[46] in the context of usual urban scaling for Eastern and Western European cities.

Discussion

We have investigated the effects of commuting networks on the GDP of urban areas using data from the US and Brazil. Inspired by the economic theory of production functions, we have modeled urban GDP as the output of a production model where population and incoming commuters (one of the simplest features of these networks) are the input variables. These generalized models assume that cities generate wealth by processes that combine their own population and the population of workers attracted from neighboring cities. We have demonstrated that these models significantly refine the description of urban GDP compared with urban scaling theory. This result allowed us to conclude that the deviations observed in the urban scaling are not random variations; instead, we have found that incoming commuters contribute to these deviations and ultimately to the GDP of urban areas.

Beyond yielding more accurate predictions for the GDP of urban areas in the US and Brazil, our modeling approach indicates that population and incoming commuters have an interconnected role in urban wealth production. In contrast to usual urban scaling, our results indicate that changes in urban GDP associated with proportionate changes in population are not independent of the initial population size. Instead, our approach suggests that the GDP of cities with different population sizes and incoming commuters respond differently to proportionate changes in these quantities. Intriguingly, increasing returns to scale of urban GDP is only possible when the product between population and number of incoming commuters exceeds a well-defined threshold. We have found that approximately 63% of all urban areas in the US and Brazil satisfy the criterion for having increasing returns to scale of GDP. For all remaining 37% urban areas of both countries, doubling population and number of incoming commuters implies less than doubling their urban GDP. Furthermore, we observe that the GDP of cities with similar population sizes or similar incoming commuters may respond quite differently to proportionate changes in population and incoming commuters.

When comparing the results between the US and Brazil, we have found that the number of urban areas with high elasticity of scale of GDP is significantly greater in Brazil, making the average elasticity of Brazilian urban areas larger than those in the US. Thus, while Brazilian cities are almost always poorer than the US ones, they usually respond more sharply to proportionate changes in population and incoming commuters. This result agrees with previous findings on urban scaling, which have shown that the GDP of urban areas in rich countries[33] or groups of rich cities[46] usually responds less sharply to population changes than their poorer counterparts. However, our results indicate that the elasticity of scale increases with the GDP of Brazilian municipalities and US counties. Thus, the negative association observed between the usual urban scaling exponent and the GDP of countries is likely to represent an aggregation effect of the elasticity of scale of individual urban areas. Moreover, we have found that Brazilian urban areas with GDP comparable to the US ones have significantly larger elasticity of scale. Large values of elasticity of scale are thus more likely to be associated with a lower degree of development than the overall wealth of urban areas.

Finally, we remark that the patterns observed in our study emerge from the interactions between population and number of commuters and by putting forward the idea that cities are not isolated entities. The significantly improved description of urban GDP in terms of population and incoming commuters corroborates this hypothesis and invites more complex approaches to describe urban indicators not only in terms of intra-city processes but through a theory that accounts for inter-city processes and interactions among cities. Other features of these commuting networks could be eventually helpful in describing the GDP or other urban indicators, and phenomenological-like approaches as the ones used here could also be used to identify variables related to city interactions. However, commuting networks are only one facet of city interactions and many other aspects such as the flow of goods, capital, and knowledge among cities are also likely to affect the economic performance of cities. Future investigations can also account for other city indicators affecting urban GDP (such as number of unemployed, education and infrastructure metrics) by generalizing our models to include new variables as well as to understand their interplay. Moreover, recent works have demonstrated that urban scaling is a time-varying phenomenon[47–49] as well as the interaction networks among cities. Despite the limitations in finding dynamic data about commuting networks, it would be interesting to extend our approach to account for dynamical features of city interactions or to use more complex artificial intelligence methods capable of learning and extracting features from commuting networks[50,51]. It is also worth remarking that all results presented here are based on administrative definitions of city (counties for US and municipalities for Brazil), and although obtaining commuting or other network-like data for several definitions of city is challenging, it would also be interesting to investigate how the exponents associated with the Cobb-Douglas and translog models are affected by these different city definitions. Regardless of the city definition and even if we manage to find ideal city boundaries, fact is that cities are not isolated entities. Thus, we hope our work stimulates further research on a theory of cities as complex interconnected systems that is far from equilibrium and depends not only on intra-city networks but also on the environment and connections with the outside world.

Methods

Data sets. *Commuting networks.* These data sets comprise the number of workers traveling from their residential urban area v_i (a US county or a Brazilian municipality) to another workplace urban area v_j. The commuting network $G(V, W)$ is thus the set of urban areas V and directed weighted edges W among them. Each

urban area $v_i \in V$ is a node in this network representation and the origin-destination flow is represented by a directed and weighted link $w_{ij} \in W$ starting from the residence location (v_i) to the workplace location (v_j). The edge weight w_{ij} represents the sum of all commuters moving daily from residence v_i to workplace v_j location. The US data set includes information about the commuting flow among the 3,220 US counties in 2015. We have obtained this data set from the US Census Bureau (USCB)[52] which in turn are produced by the American Community Survey (5-Year ACS, 2011-2015). The data set from Brazil includes information about the commuting flow between the 5,565 Brazilian municipalities in the year 2010, and we have obtained it from the Brazilian Institute of Geography and Statistics (IBGE)[53].

Population and GDP data. We have also collected the population and the urban gross domestic product (GDP) for all US counties from the National Historical Geographic Information System (NHGIS, 5-Year ACS, 2011-2015)[54]. For the Brazilian municipalities, we have obtained the analogous information for the year 2010 from the Department of Data Processing of the Brazilian Public Healthcare System (DATASUS)[55] which are in turn produced by the Brazilian Institute of Geography and Statistics (IBGE).

Ridge regression approach. While ordinary least-squares is the most common approach for estimating parameters of linear models, this method assumes predictor variables ($\log N$, $\log S$, and $\log N \log S$ in our case) to be uncorrelated. If this assumption does not hold (a problem known as multicollinearity), ordinary least-squares estimates of linear coefficients can be unstable against minor changes in data and display large standard errors. Figure 3a shows that our predictor variables are correlated and to account for this multicollinearity, we have fitted the Cobb-Douglas (Eq. 2) and translog (Eq. 3) models to our data using the ridge regression method[44]. This approach accounts for correlations among the predictor variables by finding the optimal linear coefficients (β_N, β_S, and β_c) that minimize the residual sum of squares plus a penalty term proportional to the sum of the squares of the linear coefficients, where λ is the constant of proportionality. We have estimated the optimal value of $\lambda = \lambda^*$ by minimizing the mean squared error with the leave-one-out cross-validation method. For this procedure, we have first standardized all predictors variables before finding λ^* to ensure that the regularization term is uniformly applied to all predictors. This regularization term reduces the variance and provides more reliable and stable estimates of the model parameters[44]. We have used the implementation of the ridge regression method available in the Python module `scikit-learn`[56].

References

1. Hong, I., Frank, M. R., Rahwan, I., Jung, W.-S. & Youn, H. The universal pathway to innovative urban economies. *Sci. Adv.* **6**, eaba4934 (2020).
2. Glaeser, E. L. *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier* (Penguin, New York, 2012).
3. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. USA* **104**, 7301–7306 (2007).
4. Nomaler, Ö., Frenken, K. & Heimeriks, G. On scaling of scientific knowledge production in US metropolitan areas. *PLoS ONE* **9**, e110805 (2014).
5. West, G. B. *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies* (Penguin, New York, 2017).
6. Ribeiro, H. V., Sunahara, A. S., Sutton, J., Perc, M. & Hanley, Q. S. City size and the spreading of COVID-19 in Brazil. *PLoS ONE* **15**, e0239699 (2020).
7. Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Distance to the scaling law: A useful approach for unveiling relationships between crime and urban metrics. *PLoS ONE* **8**, e69580 (2013).
8. Alves, L. G. A., Ribeiro, H. V. & Rodrigues, F. A. Crime prediction through urban metrics and statistical learning. *Physica A* **505**, 435–443 (2018).
9. Alves, L. G. A., Mendes, R. S., Lenzi, E. K. & Ribeiro, H. V. Scale-adjusted metrics for predicting the evolution of urban indicators and quantifying the performance of cities. *PLoS ONE* **10**, e0134862 (2015).
10. Sampson, R. J., Raudenbush, S. W. & Earls, F. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* **277**, 918–924 (1997).
11. Louf, R. & Barthelemy, M. Patterns of residential segregation. *PLoS ONE* **11**, e0157476 (2016).
12. Barthelemy, M. The statistical physics of cities. *Nat. Rev. Phys.* **1**, 406–415 (2019).
13. Ribeiro, H. V., Rybski, D. & Kropp, J. P. Effects of changing population or density on urban carbon dioxide emissions. *Nat. Commun.* **10**, 1–9 (2019).
14. Batty, M. The size, scale, and shape of cities. *Science* **319**, 769–771 (2008).
15. Batty, M. *The New Science of Cities* (MIT Press, Cambridge, 2013).
16. Lobo, J. *et al.* Urban science: Integrated theory from the first cities to sustainable metropolises. SSRN Scholarly Paper ID 3526940, Social Science Research Network, Rochester, NY (2020). https://papers.ssrn.com/abstract=3526940.
17. Sveikauskas, L. The productivity of cities. *Q. J. Econ.* **89**, 393–413 (1975).
18. West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126 (1997).
19. Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
20. Arbesman, S., Kleinberg, J. M. & Strogatz, S. H. Superlinear scaling for innovation in cities. *Phys. Rev. E* **79**, 016115 (2009).
21. Pan, W., Ghoshal, G., Krumme, C., Cebrian, M. & Pentland, A. Urban characteristics attributable to density-driven tie formation. *Nat. Commun.* **4**, 1–7 (2013).
22. Schläpfer, M. *et al.* The scaling of human interactions with city size. *J. R. Soc. Interface* **11**, 20130789 (2014).
23. Gonzalez, M. C., Hidalgo, C. A. & Barabasi, A.-L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).

24. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462–465 (2006).
25. Masucci, A. P., Serras, J., Johansson, A. & Batty, M. Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Phys. Rev. E* **88**, 022812 (2013).
26. Louf, R. & Barthelemy, M. Modeling the polycentric transition of cities. *Phys. Rev. Lett.* **111**, 198702 (2013).
27. Ren, Y., Ercsey-Ravasz, M., Wang, P., González, M. C. & Toroczkai, Z. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. *Nat. Commun.* **5**, 1–9 (2014).
28. Spadon, G., de Carvalho, A. . C. . P. . L. . F., ., Rodrigues-Jr, J. . F. & Alves, L. . G. . A. Reconstructing commuters network using machine learning and urban indicators. *Sci. Rep.* **9**, 1–13 (2019).
29. Keuschnigg, M., Mutgan, S. & Hedström, P. Urban scaling and the regional divide. *Sci. Adv.* **5**, eaav0042 (2019).
30. Bettencourt, L. M. A. & Zünd, D. Demography and the emergence of universal patterns in urban systems. *Nat. Commun.* **11**, 4584 (2020).
31. Altmann, E. G. Spatial interactions in urban scaling laws. *PLoS ONE* **15**, e0243390 (2020).
32. Guimera, R., Uzzi, B., Spiro, J. & Amaral, L. A. N. Team assembly mechanisms determine collaboration network structure and team performance. *Science* **308**, 697–702 (2005).
33. Ribeiro, H. V., Oehlers, M., Moreno-Monroy, A. I., Kropp, J. P. & Rybski, D. Association between population distribution and urban GDP scaling. *PLoS ONE* **16**, e0245771 (2021).
34. Heathfield, D. F. & Wibe, S. *An Introduction to Cost and Production Functions* (Macmillan, 1987).
35. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
36. Jordahl, K. Geopandas: Python tools for geographic data. *URL: https://github. com/geopandas/geopandas* (2014).
37. Moura, D. C. 3D density histograms for criteria-driven edge bundling. arXiv preprint arXiv:1504.02687 (2015).
38. Arcaute, E. *et al.* Constructing cities, deconstructing scaling laws. *J. R. Soc. Interface* **12**, 20140745 (2015).
39. Alvioli, M. Administrative boundaries and urban areas in italy: A perspective from scaling laws. *Landsc. Urban Plan.* **204**, 103906 (2020).
40. Alvioli, M. Comparative study of delineation of urban areas using imperviousness products and open data. In *Proceedings of the Geomorphometry 2020 Conference* (IRPI CNR, 2020). https://doi.org/10.30437/GEOMORPHOMETRY2020_1.
41. Newman, M. *Networks: An Introduction* (Oxford University Press, New York, 2010).
42. Cobb, C. W. & Douglas, P. H. A theory of production. *Am. Econ. Rev.* **18**, 139–165 (1928).
43. Hoerl, A. E. & Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
44. Friedman, J., Hastie, T. & Tibshirani, R. *The Elements of Statistical Learning* (Springer, 2001).
45. Christensen, L. R., Jorgenson, D. W. & Lau, L. J. Transcendental logarithmic production frontiers. *Rev. Econ. Stat.* **55**, 28–45 (1973).
46. Strano, E. & Sood, V. Rich and poor cities in Europe. An urban scaling approach to mapping the European economic transition. *PLoS ONE* **11**, e0159465 (2016).
47. Meirelles, J., Neto, C. R., Ferreira, F. F., Ribeiro, F. L. & Binder, C. R. Evolution of urban scaling: Evidence from Brazil. *PloS ONE* **13**, e0204574 (2018).
48. Bettencourt, L. M. *et al.* The interpretation of urban scaling analysis in time. *J. R. Soc. Interface* **17**, 20190846 (2020).
49. Xu, G. *et al.* Scaling laws in intra-urban systems and over time at the district level in Shanghai, China. *Physica A* **560**, 125162 (2020).
50. Liu, Z. *et al.* Learning geo-contextual embeddings for commuting flow prediction. *Proc. AAAI Conf. Artif. Intell.* **34**, 808–816 (2020).
51. Spadon, G. *et al.* Pay attention to evolution: Time series forecasting with deep graph-evolution learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021).
52. US Census Bureau. 2011-2015 5-Year ACS Commuting Flows. https://www.census.gov/data/tables/2015/demo/metro-micro/commuting-flows-2015.html (2015). Accessed on 2021-05-01.
53. Brazilian Institute of Geography and Statistics (IBGE). Census micro-data 2010. ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Resultados_Gerais_da_Amostra/Microdados (2010). Accessed on 2021-05-01.
54. Manson, S., Schroeder, J., Van Riper, D., Ruggles, S. *et al. IPUMS National Historical Geographic Information System: Version 12.0 [Database]* 39 (University of Minnesota, 2017).
55. DATASUS. *Demographic and Socioeconomic data.* http://www2.datasus.gov.br/DATASUS/index.php?area=0206 (2010). Accessed on 2021-05-01.
56. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Author contributions

L.G.A.A., D.R., and H.V.R. designed research, performed research, analyzed data, and wrote the paper.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-02327-7.

Correspondence and requests for materials should be addressed to D.R. or H.V.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.