Crime Science

More crime in cities? On the scaling laws of crime and the inadequacy of per capita rankings—a cross-country study

Marcos Oliveira[1,2]*

Abstract

Crime rates per capita are used virtually everywhere to rank and compare cities. However, their usage relies on a strong linear assumption that crime increases at the same pace as the number of people in a region. In this paper, we demonstrate that using per capita rates to rank cities can produce substantially different rankings from rankings adjusted for population size. We analyze the population–crime relationship in cities across 12 countries and assess the impact of per capita measurements on crime analyses, depending on the type of offense. In most countries, we find that theft increases superlinearly with population size, whereas burglary increases linearly. Our results reveal that per capita rankings can differ from population-adjusted rankings such that they disagree in approximately half of the top 10 most dangerous cities in the data analyzed here. Hence, we advise caution when using crime rates per capita to rank cities and recommend evaluating the linear plausibility before doing so.

Keywords: Crime rate, Population size, City, Urban scaling, Ranking, Complex systems

Introduction

In criminology, it is generally accepted that crime occurs more often in more populated regions. In one of the first works of modern criminology, Balbi and Guerry examined the crime distribution across France in 1825, revealing that some areas experienced more crime than others (Balbi and Guerry, 1829; Friendly, 2007). To compare these areas, they realized the need to adjust for population size and analyzed crime rates instead of raw numbers. This method eliminates the linear effect of population size on crime numbers and has been used to measure crime and compare cities almost everywhere—from academia to news outlets (Hall, 2016; Park and Katz, 2016; Siegel, 2011). However, this approach overlooks the potential nonlinear effects of population and, more importantly, exposes our limited understanding of the population–crime relationship.

Though different criminology theories expect a relationship between population size and crime, they tend to disagree on how crime increases with population (Chamlin and Cochran, 2004; Rotolo and Tittle, 2006). These theories predict divergent population effects, such as linear and superlinear crime growth. Despite these theoretical disputes, however, crime rates per capita are broadly used by assuming that crime increases linearly with the number of people in a region. Crucially, crime rates are often deemed to be a standard means of comparing crime in cities.

Yet the widespread adoption of crime rates is arguably due more to tradition (Boivin, 2013) rather than its ability to remove the effects of population size. Many urban indicators, including crime, have already been shown to increase nonlinearly with population size (Bettencourt et al., 2007). When we violate the linear assumption and use rates, we deal with quantities that still have

*Correspondence: M.A.Oliveira@exeter.ac.uk
[1] University of Exeter College of Engineering Mathematics and Physical Sciences, Exeter, UK
Full list of author information is available at the end of the article

population effects, thus introducing an artifactual bias into rankings and analyses.

Despite this inadequacy, we only have a limited understanding of the impact of nonlinearity on crime rates. Although previous works have investigated population–crime relationships extensively (Alves et al., 2013a; Bettencourt et al. 2010; Chang et al. 2019; Gomez-Lievano et al., 2012; Hanley et al., 2016; Yang et al., 2019), they have failed to quantify the impact of nonlinear relationships on rankings and restricted their analyses to either specific offenses or countries. The lack of comprehensive systematic studies has limited our knowledge on how the linear assumption influences crime analyses and, more critically, has prevented us from better understanding the effect of population on crime.

In this work, we analyze burglaries and thefts in 12 countries and investigate how crime rates per capita can misrepresent cities in rankings. Instead of assuming that the population–crime relationship is linear, we estimate this relationship from data using probabilistic scaling analysis (Leitão et al., 2016). We use our estimates to rank cities while adjusting for population size, and we then examine how these rankings differ from rankings based on rates per capita. In our results, we find that the linear assumption is unjustified. We show that using crime rates to rank cities can lead to rankings that considerably differ from rankings adjusted for population size. Finally, our results reveal contrasting growths of burglaries and thefts with population size, implying that different crime dynamics can produce distinct features at the city level. Our work sheds light on the population–crime relationship and suggests caution in using crime rates per capita.

Crime and population size
Different theoretical perspectives predict the emergence of a relationship between population size and crime. Three main criminology theories expect this relationship: structural, social control, and subcultural (Chamlin and Cochran, 2004; Rotolo and Tittle, 2006). In general, these perspectives agree that variations in the number of people in a region have an impact on the way people interact with one another. These theories, however, differ in the types of changes in social interaction and how they can produce a population–crime relationship.

From a structural perspective, a higher number of people increases the chances of social interaction, which increases the occurrence of crime. Two distinct rationales can explain such an increase. Mayhew and Levinger (1976) posit that crime is a product of human contact: more interaction leads to higher chances of individuals being exploited, offended, or harmed. They claim that a

larger population size raises the number of opportunities for interaction at an increasing rate, which would lead to a superlinear crime growth with population size (Chamlin and Cochran, 2004). In contrast, Blau (1977) implies a linear population–crime relationship. He posits that population aggregation reduces spatial distance among individuals, thereby promoting different social associations such as victimization. At the same time, as conflictive association increases, other integrative associations also increase, leading to a linear growth of crime (Chamlin and Cochran, 2004). Notably, the structural perspective focuses on the quantitative consequences of population growth.

The social control perspective advocates that changes in population size have a qualitative impact on social relations, which weakens informal social control mechanisms that inhibit crime (Groff, 2015). From this perspective, crime relates to two aspect of a population: size and stability. A larger population size leads to higher population density and heterogeneity—not only do individuals have more opportunities for social contacts, but they are also often surrounded by strangers (Wirth, 1938). This situation makes social integration difficult and promotes a high anonymity, which encourages criminal impulses and harms a community's ability to socially constrain misbehavior (Freudenburg, 1986; Sampson, 1986). Similarly, from a systemic viewpoint, any change (i.e., increase or decrease) in population size can have an impact on crime numbers (Rotolo and Tittle, 2006). From this viewpoint, the understanding is that regular and sustained social interactions produce community networks with effective mechanisms of social control (Bursik and Webb, 1982). Population instability, however, hinders the construction of such networks. In communities with unstable population size, residents avoid socially investing in their neighborhoods, which hurts community organization and weakens social control, thus increasing misbehavior and crime (Miethe et al., 1991; Sampson, 1988).

Both social-control and structural perspectives solely focus on individuals' interactions without considering their private interests. These perspectives pay little attention to how unconventional interests increase with urbanization (Fischer, 1975) and how these interests relate to misbehavior.

In contrast, the subcultural perspective advocates that population concentration brings together individuals with shared interests, which produces private social networks built around these interests, thereby promoting social support for behavioral choices. Fischer (1975) posits that population size has an impact on the creation, diffusion, and intensification of unconventional interests. He proposes

that large populations have a sufficient number of people with specific shared interests, thus enabling social interaction and lead to the emergence of subcultures. The social networks surrounding a subculture bring normative expectations that increase the likelihood of misbehavior and crime (Fischer, 1975,, 1995).

These three perspectives—structural, social control, and subcultural—expect that a higher number of people in an area leads to more crime in that area. In the case of cities, we know that population size is indeed a strong predictor of crime (Bettencourt et al., 2007) . The existence of a population–crime relationship implies that we must adjust for population size to analyze crime in cities properly.

Crime rates per capita

In the literature, the typical solution for removing the effect of population size from crime numbers is to use ratios such as the following:

$$\text{crime rate per capita} = \frac{\text{crime}}{\text{population}}. \tag{1}$$

This ratio is often used together with a multiplier that contextualizes the quantity (e.g., crime per 100,000 inhabitants; Boivin, 2013). However, even though crime rates are popularly used, they present at least two inadequacies. First, the way in which we define population affects crime rates. The common approach is to use resident population (e.g., census data) to estimate rates, but this practice can distort the picture of crime in a place: crime is not limited to residents (Gibbs and Erickson, 1976), and cities attract a substantial number of non-residents (Stults and Hasbrouck, 2015). Instead, researchers suggest using ambient population (Andresen, 2006,, 2011) and accounting for criminal opportunities, which depends on the type of crime (Boggs, 1965; Clarke, 1984; Cohen et al., 1985; Harries, 1981).

Second, Eq. (1) assumes that the population–crime relationship is linear. The rationale behind this equation is that we have a relationship of the form

$$\text{crime} \sim \text{population}, \tag{2}$$

which means that crime can be linearly approximated via population. Given this linear assumption, when we divide crime by population in Eq. (1), we are trying to cancel out the effect of population on crime. This assumption implies that crime increases at the same pace as population growth. However, not all theoretical perspectives agree with this type of growth, and many urban indicators, including crime, have been shown to increase with population size in a nonlinear fashion (Bettencourt et al., 2007).

Cities and scaling laws

Much research has been devoted to understanding urban growth and its impact on indicators such as gross domestic product, total wages, electrical consumption, and crime (Bettencourt et al., 2007, 2010; Bettencourt, 2013; Gomez-Lievano et al., 2016). Bettencourt et al. (2007) have found that a city's population size, denoted by N, is a strong predictor of its urban indicators, denoted by Y, exhibiting the following relationship:

$$Y \sim N^{\beta}. \tag{3}$$

This so-called scaling law tells us that, given the size of a city, we expect certain levels of wealth creation, knowledge production, criminality, and other urban aspects. This expectation suggests general processes underlying urban development (Bettencourt et al., 2013) and indicates that regularities exist in cities despite their idiosyncrasies (Oliveira and Menezes, 2019). To understand this scaling and the urban processes better, we can examine the exponent β, which describes how an urban indicator grows with population size.

Bettencourt et al. (2007) presented evidence that different categories of urban indicators exhibit distinct growth regimes. They showed that social indicators grow faster than infrastructural ones (see Fig. 1A). Specifically, social indicators, such as the number of patents and total wages, increase superlinearly with population size (i.e., $\beta > 1$), meaning that these indicators grow at an increasing rate with population. In the case of infrastructural aspects (e.g., road surface, length of electrical cables), an economy of scale exists. As cities grow in population size, these urban indicators increase at a slower pace with $\beta < 1$ (i.e., sublinearly). In both scenarios, because of nonlinearity, we should be careful with per capita analyses.

When we violate the linear assumption of per capita ratios, we deal with quantities that can misrepresent an urban indicator. To demonstrate this, we use Eq. (3) to define the per capita rate C of an urban indicator as follows:

$$C = \frac{Y}{N} \sim N^{\beta-1},$$

which implies that rates are independent from population only when β equals to one—when $\beta \neq 1$, population is not cancelled out from the equation. In these nonlinear cases, per capita rates can inflate or deflate the representation of an urban indicator depending on β (see Fig. 1B). This misrepresentation occurs because population still has an effect on rates. By definition, we expect that per capita rates are higher in larger cities when
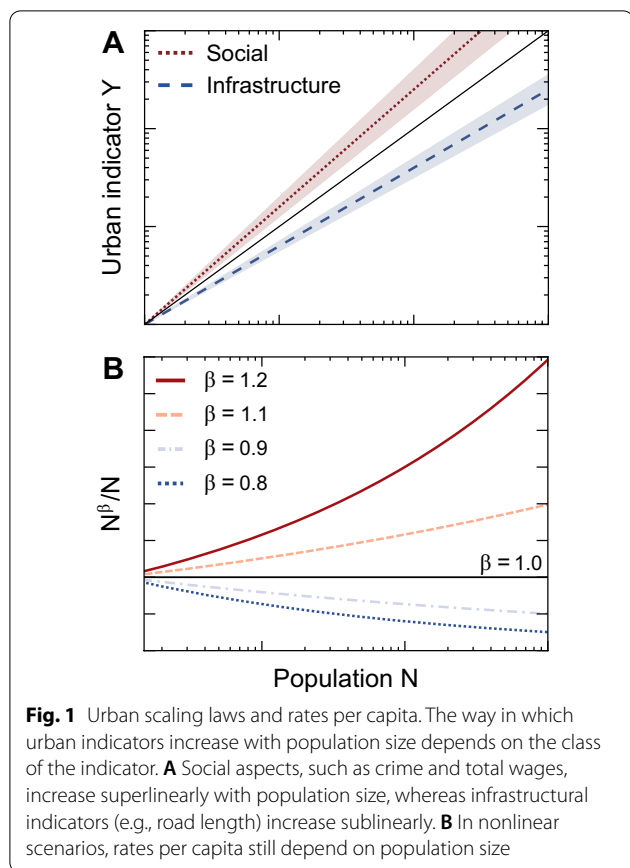
Fig. 1 Urban scaling laws and rates per capita. The way in which urban indicators increase with population size depends on the class of the indicator. **A** Social aspects, such as crime and total wages, increase superlinearly with population size, whereas infrastructural indicators (e.g., road length) increase sublinearly. **B** In nonlinear scenarios, rates per capita still depend on population size

Yang et al., 2019). Remarkably, the existence of these scaling laws of crime suggests fundamental urban processes that relate to crime, independent of cities' particularities.

This regularity manifests itself in the so-called scale-invariance property of scaling laws. It is possible to show that Eq. (3) holds the following property:

$$Y(\kappa N) = g(\kappa)Y(N), \tag{4}$$

where $g(\kappa)$ does not depend on N (Thurner et al., 2018). From a modeling perspective, this relationship reveals two aspects about crime. First, we can predict crime numbers in cities via a populational scale transformation κ (Bettencourt et al., 2013). This transformation is independent of population size but depends on β, which tunes the relative increase in crime such that $g(\kappa) = \kappa^{\beta}$. Second, Eq. (4) implies that crime is present in any city, independent of size. This implication arguably relates to the Durkheimian concept of crime normalcy in that crime is seen as a normal and necessary phenomenon in societies, provided that its numbers are not unusually high (Durkheim, 1895). Broadly speaking, the scale-invariance property tells us that crime in cities is associated with population in a somewhat predictable fashion. Crucially, this property might give the impression that such regularity is independent of crime type.

However, different types of crime are connected to social mechanisms in different ways (Hipp and Steenbeek, 2016) and exhibit unique temporal (Miethe et al., 2005; Oliveira et al., 2018) and spatial characteristics (Andresen and Linning, 2012; Oliveira et al., 2015, 2017; White et al., 2014). It is plausible that the scaling laws of crime depend on crime type. Nevertheless, the literature has mostly focused on either specific countries or crime types. Few studies have systematically examined the scaling of different crime types, and the focus on specific countries has prevented us from better understanding the impact of population on crime. Likewise, the lack of a comprehensive systematic study has limited our knowledge about the impact of the linear assumption on crime rates. We still fail to understand how per capita analyses can misrepresent cities in nonlinear scenarios.

In this work, we characterize the scaling laws of burglary and theft in 12 countries and investigate how crime rates per capita can misrepresent cities in rankings. Instead of assuming that the population−crime relationship is linear, as described in Eq. (2), we investigate this relationship under its functional form as follows:

$$\text{crime} \sim f(\text{population}).$$

$\beta > 1$, whereas when $\beta < 1$, we expect larger cities to have lower rates. When we use rates to compare cities in nonlinear situations, we introduce an artifactual bias. To compare cities properly, previous works have proposed scaled-adjusted indicators that account for population size (Alves et al., 2013a; Bettencourt et al., 2010), supporting the need for population adjustment but failing to quantify the impact of the linear assumption on rankings of urban indicators.

More crime in cities?

In the case of crime, researchers have found a superlinear growth with population size. Bettencourt et al. (2007) showed that serious crime in the United States exhibits superlinear scaling with exponent $\beta \approx 1.16$, and some evidence has confirmed similar superlinearity for homicides in Brazil, Colombia, and Mexico (Alves et al., 2013b; Gomez-Lievano et al., 2012). Previous works have also found that different kinds of crime in the United Kingdom and in the United States present nonlinear scaling relationships (Chang et al., 2019; Hanley et al., 2016;

Specifically, we examine the plausibility of scaling laws to describe the population–crime relationship. To estimate the scaling laws, we use probabilistic scaling analysis, which enables us to characterize the scaling laws of crime. We use our estimates to rank cities while accounting for the effects of population size. Finally, we compare these adjusted rankings with rankings based on per-capita rates (i.e., with the linear assumption).

Results

We use data from 12 countries to investigate the relationship between population size and crime at the city level (see the appendix for data sources). Specifically, we examine annual data from Belgium, Canada, Colombia, Denmark, France, Italy, Mexico, Portugal, South Africa, Spain, the United Kingdom, and the United States (see Table 1). In this work, we characterize how crime increases with population size in each country, focusing on burglary and theft. We analyze both crimes in all considered countries, except Mexico, Portugal, and Spain, where we only have data for one type of offense.

The scaling laws of crime in cities

To assess the relationship between crime Y and population size N (see Fig. 2), we model P(Y|N) using probabilistic scaling analysis (see the Methods section). In our study, we examine whether this relationship follows the general form of $Y \sim N^{\beta}$. First, we estimate β from data, and we then evaluate the plausibility of the model ($p > 0.05$) and the evidence for nonlinearity (i.e., $\beta \neq 1$). Our results reveal that Y and N often exhibit a nonlinear relationship, depending on the type of offense.

In most of the considered countries, theft increases with population size superlinearly, whereas burglary tends to increase linearly (see Fig. 3). Precisely, in 9 out of 11 countries, we find that β for theft is above one; our results indicate linearity for theft (i.e., absence of nonlinear plausibility) in Canada and South Africa. In the case of burglary, we are unable to reject linearity in 7 out of 10 countries; in France and the United Kingdom, we find superlinearity, and in Canada, sublinearity. In almost all considered data sets, these estimates are consistent over two consecutive years in the countries for which we have data for different years (see Appendix 1).

Our results suggest that the general form of $Y \sim N^{\beta}$ is plausible in most countries, but that this compatibility depends on the offense. We find that burglary data are compatible with the model ($p > 0.05$) in 80% of the considered countries. In the case of theft, the superlinear models are compatible with data in five out of nine countries. We note that in Canada and South Africa, where we are unable to reject linearity for theft, the linear model also lacks compatibility with data.

We find that the estimates of β for each offense often have different values across countries—for example, the superlinear estimates of β for theft range from 1.10 to 1.67. However, when we analyze each country separately, we find that β for theft tends to be larger than β for burglary in each country, except for France and the United Kingdom.

In summary, we find evidence for a nonlinear relationship between crime and population size in more than half of the considered data sets. Our results indicate that crime often increases with population size at a pace that is different from per capita. This relationship

Table 1 Burglary and theft annual statistics in 12 countries

Country	n	Theft			Burglary		
		$\bar{y}$	S	y_{max}	$\bar{y}$	S	y_{max}
Belgium	588	60.84	286.51	4397	95.60	209.02	2721
Canada	283	1115.14	3393.88	37,150	293.90	791.13	7782
Colombia	513	182.04	1514.68	36,306	40.08	228.06	4856
Denmark	98	1157.67	3851.29	38,011	330.71	330.60	2157
France	100	8311.12	12,400.34	108,846	2389.94	2515.24	12,511
Italy	107	17,470.72	30,860.27	218,052	2217.50	2642.61	18,101
Mexico	1659	237.56	959.59	14,999	–	–	–
Portugal	279	–	–	–	51.38	86.91	850
South Africa	199	2305.23	8758.52	93,793	1190.03	3212.93	28,143
Spain	144	7846.72	25,111.99	236,026	–	–	–
United Kingdom	313	1763.43	1965.61	19,766	620.98	685.40	4825
United States	8337	471.82	2345.27	108,376	127.33	626.16	19,859

Number of data points n, sample mean $\bar{y}$, sample standard deviation S, and maximum value y_{max}
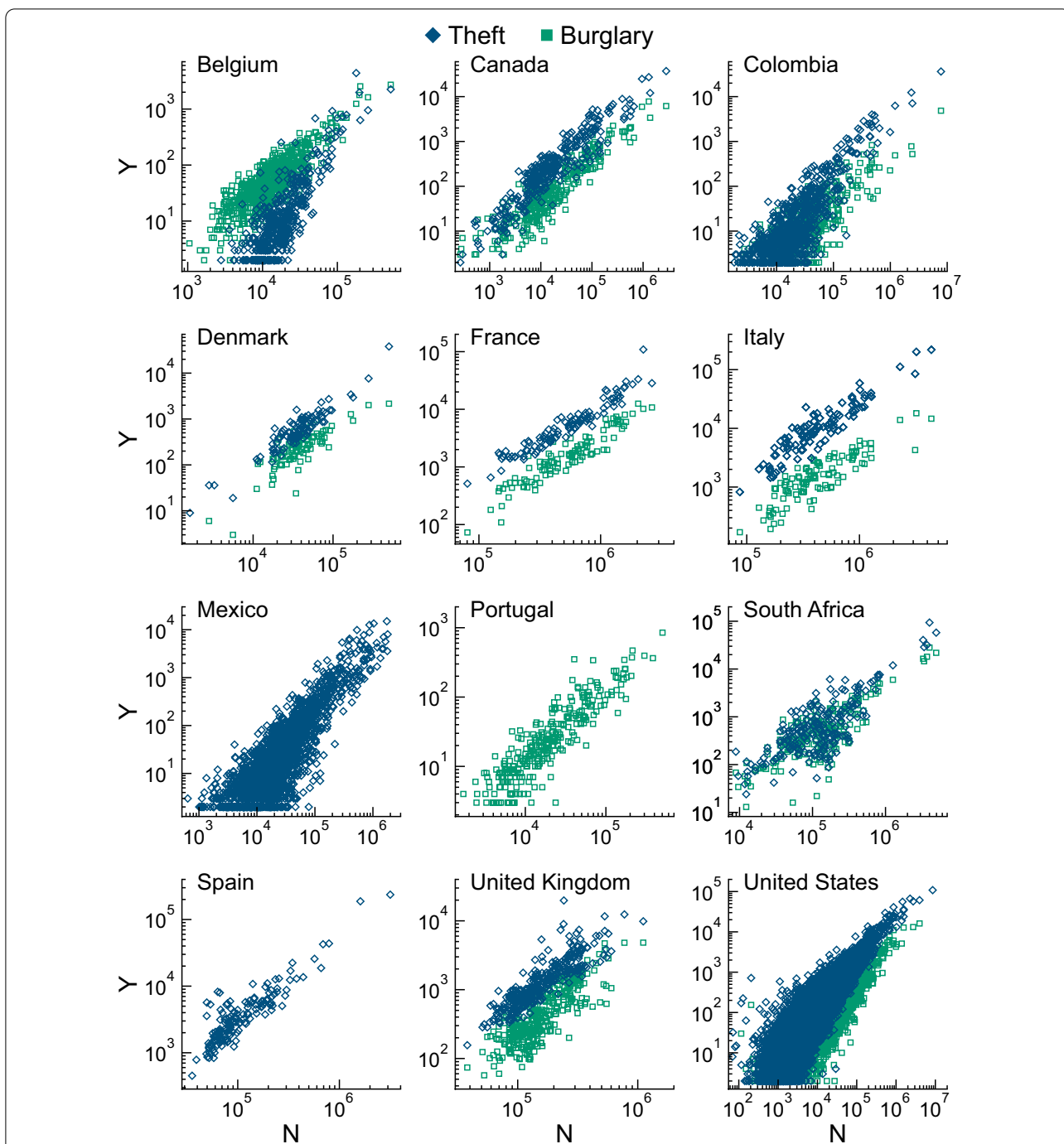
Fig. 2 The population–crime relationship in 12 countries. Different criminology theories expect a relationship between population size and crime, predicting divergent population effects, such as linear and superlinear crime growth. Despite these theoretical disputes, however, crime rates per capita are broadly used by assuming that crime increases linearly with population size

implies that analyses with a linear assumption might create distorted pictures of crime in cities. To understand such distortions, we must examine how nonlinearity influences comparisons of crime in cities, when linearity is assumed.

The inadequacy of crime rates and per capita rankings

We investigate how crime rates of the form $C = Y/N$ introduce bias in the comparisons and rankings of cities. To understand this bias, we use Eq. (3) to rewrite crime rate as $C \sim N^{\beta-1}$. This relationship implies that

Fig. 3 The scaling laws of crime. We find evidence for a nonlinear relationship between crime and population size in more than half of the data sets. In most considered countries, theft exhibits superlinearity, whereas burglary tends to display linearity. In the plot, the lines represent the error bars for the estimated β of each country–crime for two consecutive years; circles denote a lack of nonlinearity plausibility; triangles represent superlinearity, and upside-down triangles indicate sublinearity
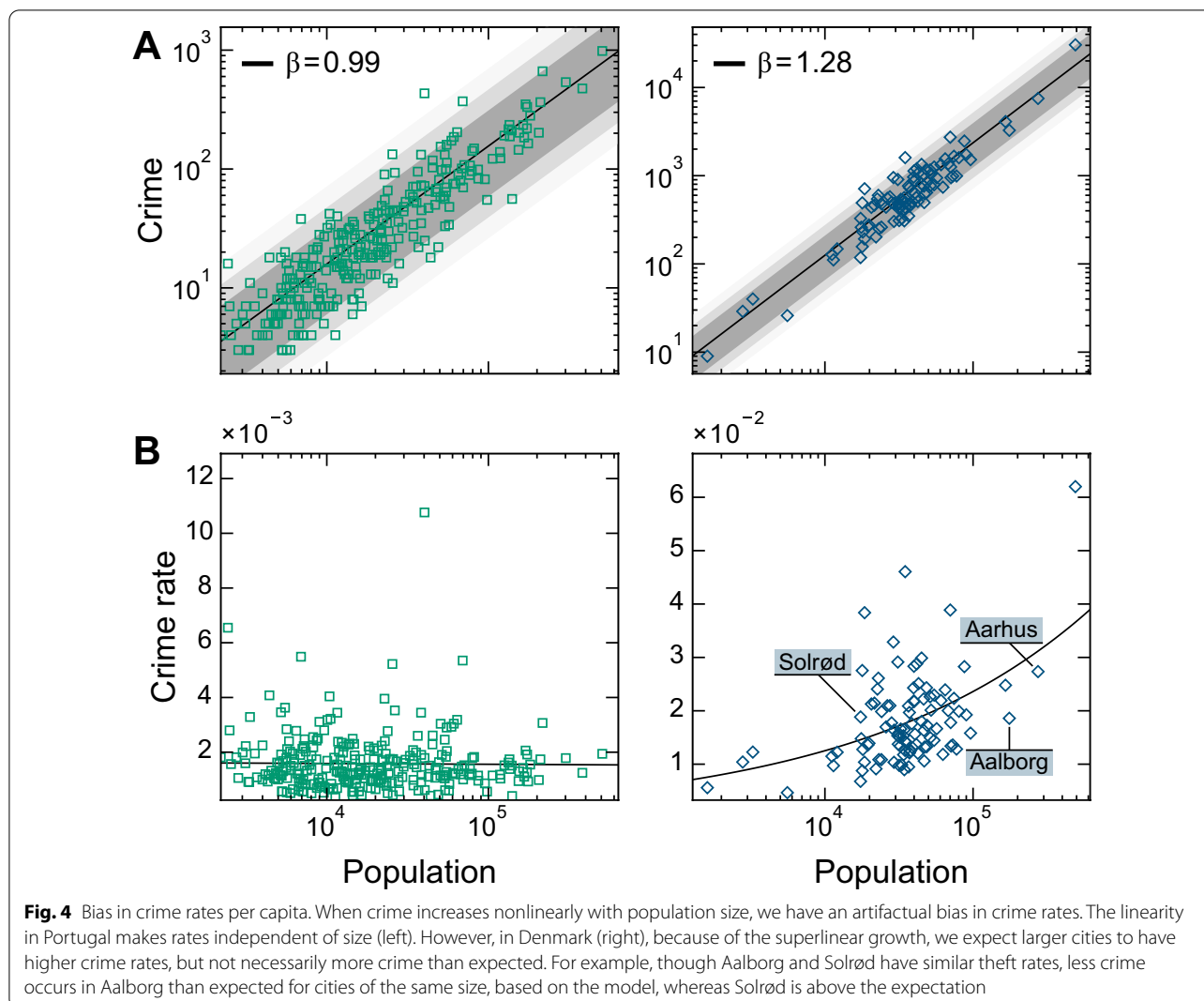
crime rate depends on population size when $\beta \neq 1$. For example, in Portugal and Denmark, this dependency is clear when we analyze burglary and theft numbers (see Fig. 4). In the case of burglary in Portugal, linearity makes C independent of population size. In Denmark, since theft increases superlinearly, we expect rates to increase with population size. In this country, based on data, the expected theft rate of a small city is lower than the rates of larger cities. We must account for this tendency in order to compare crime in cities; otherwise, we introduce bias against larger cities.

To account for the population–crime relationship found in data, we compare cities using the model $P(Y|N)$ as the baseline. We compare the number of crimes in a city with the expectation of the model. For each city i with population size n_i, we evaluate the z score of the city

with respect to $P(Y|N = n_i)$. The z score indicates how much more or less crime a particular city has in comparison to cities with a similar population size, as expected by the model. These z scores enable us to compare cities in a country and rank them while accounting for population size differences. In contrast, crime rates per capita only adjust for population size in the linear scenario. This approach is similar to previously proposed indicators that adjust for population size (Alves et al., 2013a; Bettencourt et al., 2010). In our case, the adjustment also accounts for the variance. We denote this kind of analysis as a comparison adjusted for the population–crime relationship.

For example, in Denmark, the theft rate in the municipality of Aalborg (≈ 0.0186) is almost the same as in Solrød (≈ 0.0188). However, less crime occurs in Aalborg than expected for cities of a similar size, while crime in

Fig. 4 Bias in crime rates per capita. When crime increases nonlinearly with population size, we have an artifactual bias in crime rates. The linearity in Portugal makes rates independent of size (left). However, in Denmark (right), because of the superlinear growth, we expect larger cities to have higher crime rates, but not necessarily more crime than expected. For example, though Aalborg and Solrød have similar theft rates, less crime occurs in Aalborg than expected for cities of the same size, based on the model, whereas Solrød is above the expectation

Solrød is above the model expectation (see Fig. 4B). This disagreement arises because of the different population sizes. Since Aalborg is more than 10 times larger than Solrød, we expect rates in Aalborg to be larger than in Solrød. When we account for this tendency and evaluate their z scores, we find that the z score of Aalborg is -2.47, whereas in Solrød the z score is 2.43.
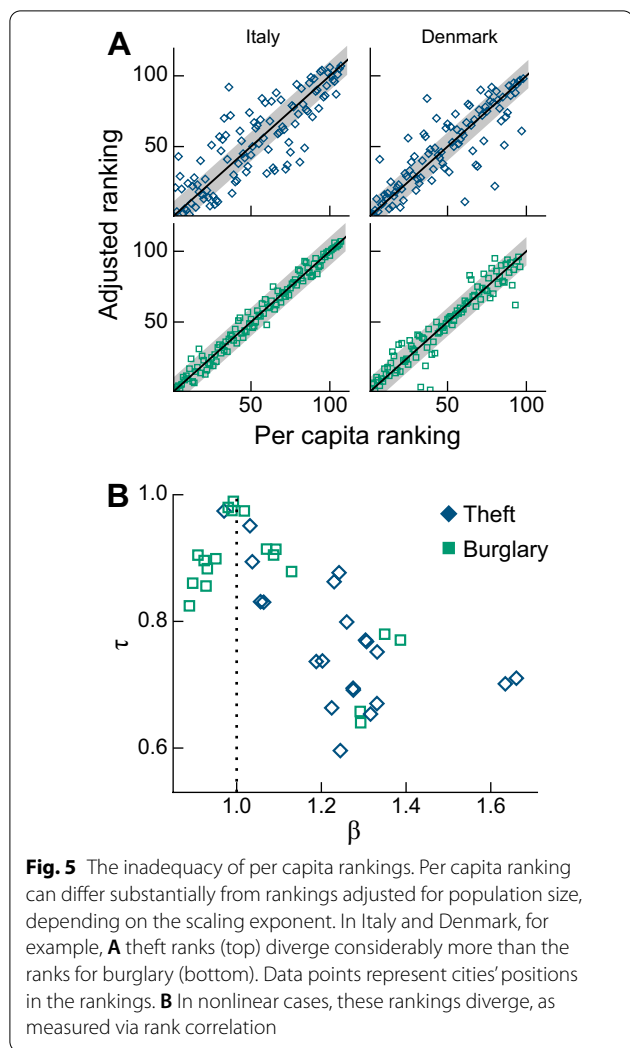
Such inconsistencies have an impact on the crime rankings of cities. The municipality of Aarhus, in Denmark, for example, is ranked among the top 12 cities with the highest theft rate in the country. However, when we account for population–crime relationship using z scores, we find that Aarhus is only at the end of the top 54 rankings.

To understand these variations systematically, we compare rankings based on crime rates with rankings that account for the population–crime relationship (i.e., adjusted rankings). Our results reveal that these

two rankings create distinct representations of cities. For each considered data set, we rank cities based on their z scores and crime rates C, and we then examine the change in the rank of each city. According to our findings, the positions of the cities can change substantially. For instance, in Italy, half of the cities have theft rate ranks that diverge in at least 11 positions from the adjusted ranking (Fig. 5A). This disagreement means that these rankings disagree for approximately half of the top 10 most dangerous cities.

We evaluate these discrepancies by using the Kendall rank correlation coefficient τ to measure the similarity between crime rates and adjusted rankings in the considered countries. We find that these rankings can differ considerably but converge when $\beta \approx 1$. The τ coefficients for the data sets range from 0.6 to 1.0, exhibiting a dependency on the type of crime; or more specifically, on the scaling (Fig. 5B). As expected, as β approaches 1,

Fig. 5 The inadequacy of per capita rankings. Per capita ranking can differ substantially from rankings adjusted for population size, depending on the scaling exponent. In Italy and Denmark, for example, **A** theft ranks (top) diverge considerably more than the ranks for burglary (bottom). Data points represent cities' positions in the rankings. **B** In nonlinear cases, these rankings diverge, as measured via rank correlation

the rankings are more similar to one another. For example, in Italy, in contrast to theft, the burglary rate ranking of half of the cities only differs from the adjusted ranking in a maximum of two positions (Fig. 5A).

Discussion and conclusion

Despite its popularity, comparing cities via crime rates without accounting for population size has a strong assumption that crime increases at the same pace as the number of people in a region. Though previous works have widely investigated the population–crime relationship, they have failed to quantify the impact of nonlinear relationships on rankings and restricted their analyses to either specific offenses or countries. In this work, we analyze crime in different countries to investigate how crime grows with population size and how the widespread assumption of linear growth influences cities' rankings.

First, we analyzed crime in cities from 12 countries to characterize the population–crime relationship statistically, examining the plausibility of scaling laws to describe this relationship. Then, we used our estimates to rank cities and compared how those rankings differ from rankings based on rates per capita.

Our results showed that the assumption of linear crime growth is unfounded. In more than half of the considered data sets, we found evidence for nonlinear crime growth—that is, crime often increases with population size at a different pace than per capita. This nonlinearity introduces a population effect into crime rates, influencing rankings. We demonstrated that using crime rates to rank cities substantially differs from ranking cities adjusted for population size.

These findings imply that using crime rates per capita—though deemed a standard measure in criminal justice statistics—can create a distorted view of cities' rankings. For example, in superlinear scenarios, we expect larger cities to have higher crime rates. In this case, when we use rates to rank cities, we build rankings whereby large cities are at the top. But, these cities might not experience more crime than what we expect from places with a similar population size. It is an artifactual bias introduced by population effects still present in crime rates.

Such effects arise from nonlinear population effects that persist in rates due to the linear assumption. This assumption is more than just a statistical subtlety. By assuming linearity, we essentially overlook cities' context: we ignore the actual impact of population size on crime and how this impact depends on crime type, country, and aggregation units, among other things. For instance, our results indicate that in thefts, linearity is an exception rather than the rule. The indiscriminate use of crime rates neglects significant population–crime interactions that should be considered in order to compare crime in cities properly.

As a result of this inadequacy, we advise caution when using crime rates per capita to compare cities. We recommend evaluating linear plausibility before comparing crime rates. In general, we suggest comparing cities via the *z* scores computed using the approach (Leitão et al., 2016) discussed in the manuscript, thereby avoiding crime rates. It is important to emphasize that this inadequacy in rates is relevant only when comparing cities of different population sizes. In analyses without comparisons, a place's crime rate can be seen as a rough indicator that contextualizes crime numbers relative to population size. Additionally, when cities have the same size, comparing crime rates boils down to comparing raw crime numbers.

Table 2 β estimates for the case of thefts using log-normal and normal fluctuations

	Log-normal		Gaussian	
	$\delta = 2$	$\delta \in [1,3]$	$\delta = 1$	$\delta \in [1,2]$
Belgium (2015)	1.63 (0.12) ↗	1.64 (0.12) ↗	2.11 (0.27) ↗	1.67 (0.17) ↗
Belgium (2016)	1.66 (0.15) ↗	1.66 (0.14) ↗	2.10 (0.18) ↗	1.75 (0.19) ↗
Canada (2015)	1.09 (0.06) ↗	1.04 (0.05) →	1.07 (0.11) →	1.04 (0.06) →
Canada (2016)	1.03 (0.04) →	1.04 (0.05) →	1.06 (0.34) ○	1.03 (0.05) →
Colombia (2013)	1.25 (0.07) ↗	1.23 (0.07) ↗*	1.89 (0.09) ↗	1.31 (0.08) ↗
Colombia (2014)	1.26 (0.07) ↗	1.24 (0.09) ↗*	1.89 (0.09) ↗	1.36 (0.08) ↗
Denmark (2015)	1.28 (0.10) ↗*	1.27 (0.13) ↗*	1.45 (0.33) ↗	1.27 (0.14) ↗
Denmark (2016)	1.27 (0.14) ↗*	1.28 (0.18) ↗*	1.58 (0.37) ↗	1.28 (0.18) ↗
France (2013)	1.24 (0.09) ↗	1.23 (0.07) ↗*	1.59 (0.44) ↗	1.30 (0.12) ↗
France (2014)	1.24 (0.10) ↗	1.22 (0.08) ↗	1.70 (0.57) ↗	1.34 (0.18) ↗
Italy (2014)	1.33 (0.11) ↗*	1.31 (0.10) ↗*	1.37 (0.15) ↗	1.31 (0.09) ↗
Italy (2015)	1.32 (0.09) ↗*	1.29 (0.11) ↗*	1.35 (0.14) ↗	1.29 (0.10) ↗
Mexico (2015)	1.30 (0.04) ↗	1.31 (0.04) ↗	1.98 (0.02) ↗	1.32 (0.04) ↗
Mexico (2016)	1.26 (0.04) ↗	1.26 (0.04) ↗	1.98 (0.01) ↗	1.30 (0.05) ↗
South Africa (2016)	0.97 (0.11) →*	0.99 (0.10) →*	1.33 (0.20) ↗	1.02 (0.11) →
Spain (2015)	1.18 (0.11) ↗	1.19 (0.11) ↗	1.27 (0.19) ↗	1.22 (0.12) ↗
Spain (2016)	1.20 (0.11) ↗	1.20 (0.11) ↗	1.31 (0.20) ↗	1.24 (0.13) ↗
United Kingdom (2015)	1.24 (0.07) ↗	1.31 (0.10) ↗	1.45 (0.30) ↗	1.55 (0.32) ↗
United Kingdom (2016)	1.26 (0.09) ↗	1.33 (0.10) ↗	1.50 (0.37) ↗	1.59 (0.35) ↗
United States (2014)	1.12 (0.01) ↗	1.06 (0.01) ↗	1.07 (0.06) ↗	1.04 (0.04) ↗
United States (2015)	1.13 (0.01) ↗	1.06 (0.01) ↗	1.08 (0.07) ↗	1.05 (0.04) ↗

Table 3 β estimates for the case of burglaries using log-normal and normal fluctuations

	Log-normal		Gaussian	
	$\delta = 2$	$\delta \in [1,3]$	$\delta = 1$	$\delta \in [1,2]$
Belgium (2015)	1.10 (0.06) ↗	1.09 (0.05) ○	1.21 (0.11) ↗	1.09 (0.05) ○
Belgium (2016)	1.08 (0.06) ○	1.07 (0.05) ○	1.18 (0.10) ↗	1.08 (0.05) ○
Canada (2015)	0.93 (0.05) ○	0.93 (0.04) ↘*	1.04 (0.10) →	0.95 (0.06) ○
Canada (2016)	0.91 (0.04) ↘*	0.90 (0.05) ↘*	1.00 (0.10) →	0.90 (0.04) ↘
Colombia (2013)	0.90 (0.07) ○*	0.93 (0.07) →	1.18 (0.44) →	0.96 (0.07) →
Colombia (2014)	0.94 (0.07) →*	0.95 (0.06) →	1.16 (0.51) →	0.99 (0.07) →
Denmark (2015)	1.11 (0.26) →	0.91 (0.14) →	0.92 (0.14) →*	0.93 (0.13) →*
Denmark (2016)	1.15 (0.24) →	0.89 (0.15) →	0.90 (0.13) →	0.92 (0.17) →
France (2013)	1.29 (0.09) ↗*	1.27 (0.09) ↗*	1.31 (0.11) ↗*	1.27 (0.09) ↗
France (2014)	1.29 (0.10) ↗*	1.27 (0.10) ↗*	1.34 (0.10) ↗	1.27 (0.09) ↗
Italy (2014)	1.13 (0.15) →*	1.11 (0.16) →*	1.09 (0.17) →	1.09 (0.12) →*
Italy (2015)	1.09 (0.15) →*	1.07 (0.13) →*	1.06 (0.15) →*	1.05 (0.12) →*
Portugal (2015)	0.99 (0.06) →*	0.98 (0.05) →*	1.13 (0.13) →	0.99 (0.10) →
Portugal (2016)	1.02 (0.05) →	1.01 (0.06) →	1.11 (0.09) ○	1.05 (0.10) →
South Africa (2016)	0.91 (0.09) →	0.91 (0.08) →	1.07 (0.09) ○	0.97 (0.12) →
United Kingdom (2015)	1.39 (0.11) ↗*	1.42 (0.10) ↗*	1.47 (0.13) ↗	1.40 (0.10) ↗
United Kingdom (2016)	1.35 (0.11) ↗*	1.36 (0.10) ↗*	1.46 (0.14) ↗	1.37 (0.11) ↗
United States (2014)	0.99 (0.01) →	0.99 (0.01) →	1.19 (0.11) ↗	1.07 (0.05) ↗
United States (2015)	0.98 (0.01) →	0.98 (0.01) ○	1.17 (0.08) ↗	1.07 (0.06) ↗

In summary, in this work, we shed light on the population–crime relationship. The linear assumption is exhausted and expired. We have resounding evidence of nonlinearity in crime, which disallows us from unjustifiably assuming linearity. In light of our results, we also note that the scaling laws are plausible models only for half of the considered data sets. Better models are thus needed—in particular, models that account for the fact that different crime types relate to population size differently. More adequate models will help us better understand the relationship between population and crime.

Limitations

Our work presents limitations related to the way in which we define population, crime, and cities. First, we note that crime rates depend on how we define population; in our study, we define it as the resident population (i.e., census data). However, crime is not limited to residents (Gibbs and Erickson, 1976), and cities attract a significant number of non-residents (Stults and Hasbrouck, 2015). We highlight that this limitation is not specific to our study, and crime rates are often measured using resident population. Previous works have suggested using ambient population and accounting for the number of targets (Andresen, 2006, 2011; Boggs, 1965). Collecting this data, however, is challenging, especially when dealing with different countries. Future research should investigate crime rates and scaling laws using other definitions of population, particularly using social media data (Malleson and Andresen, 2016; Pacheco et al., 2017).

Second, scaling analyses depend on the definition of what constitutes a city (Arcaute et al., 2014). In the literature, definitions include legal divisions (e.g., counties, municipalities) and data-driven delineations based on population density and economic interactions (Cottineau et al., 2017). It is possible that different city definitions yield divergent scaling regimes for the same urban indicator (Louf and Barthelemy, 2014). In our work, we only have access to crime data regarding specific aggregation units, and we thus define cities based on official legal divisions by using census data. City definitions in our analysis consequently depend on the country. We emphasize that we investigate whether per capita rankings are justified under a given city definition. Nevertheless, we believe that even though the use of other city definitions might change our quantitative results, our qualitative results are robust: the inadequacy of crime rates is independent of city definitions. When analyzing different definitions of cities, future research should examine scaling divergences as an opportunity to understand the population–crime relationship better.

Finally, cross-national crime analyses have methodological challenges due to international differences in crime definitions, police and court practices, and reporting rates, among other things (Takala and Aromaa, 2008). Although we avoid direct comparisons of countries' absolute crime numbers in our work, we compare their growth exponents. In this comparison, we assume that cross-national differences have a negligible impact on how crime increases with population, particularly regarding the crime types we analyzed. We understand that some offenses (e.g., sexual assault, drug trafficking) are more sensitive to cross-national comparisons than the offenses we analyzed here (Harrendorf et al., 2010; Harrendorf, 2018). Collecting high-quality international comparative data could help future works in disentangling cross-national differences.

Methods

Probabilistic scaling analysis

We use probabilistic scaling analysis to estimate the scaling laws of crime. Instead of analyzing the linear form of Eq. (3), we use the approach developed by Leitão et al. (2016) to estimate the parameters of a distribution $Y|N$ that has the following expectation:

$$E[Y|N] = \lambda N^{\beta}, \tag{5}$$

that is, N scales the expected value of an urban indicator (Bettencourt et al., 2013; Gomez-Lievano et al., 2012; Leitão et al., 2016). Note that this method does not assume that the fluctuations around $\ln y$ and $\ln x$ are normally distributed (Leitão et al., 2016). Instead, we compare models for $P(Y|N)$ that satisfy the following conditional variance:

$$V[Y|N] = \gamma E[Y|N]^{\delta}, \tag{6}$$

where typically $\delta \in [1, 2]$, since urban systems have been previously shown to exhibit non-trivial fluctuations around the mean—the so-called Taylor's law (Hanley et al., 2014). To estimate the scaling laws, we maximize the log-likelihood

$$\mathcal{L} = \ln P(y_1, \ldots, y_K | n_1, \ldots, n_K) = \sum_{i=1}^{K} \ln P(y_i | n_i),$$

since we assume y_i as an independent realization from $P(Y|N)$. In this work, we use an implementation developed by Leitão et al. (2016) that maximizes the log-likelihood with the "L-BFGS-B" algorithm. We model $P(Y|N)$ using Gaussian and log-normal distributions in order to analyze whether accounting for the size-dependent variance influences the estimation. In the case of the Gaussian, the conditions from Eq. (5) and Eq. (6) are satisfied with

$$\mu_{\mathsf{N}}(x) = \alpha x^{\beta} \quad \text{and} \quad \sigma_{\mathsf{N}}^2(x) = \gamma(\alpha x^{\beta})^{\delta},$$

whereas in the case of the log-normal distribution,

$$\mu_{\mathsf{LN}}(x) = \ln \alpha + \beta \ln x - \frac{1}{2}\sigma_{\mathsf{LN}}^2(x) \quad \text{and}$$

$$\sigma_{\mathsf{LN}}^2(x) = \ln\left[1 + \gamma(\alpha x^{\beta})^{\delta-2}\right].$$

In the log-normal case, note that, if $\delta = 2$, then the fluctuations are independent of N; thus this would be the same as using the minimum least-squares approach (Leitão et al., 2016). With this framework, we compare models that have fixed δ against models wherein δ is also included in the optimization process. In the case of the Gaussian, we have fixed $\delta = 1$ and free $\delta \in [1,2]$, whereas in the case of the log-normal, we have fixed $\delta = 2$ and free $\delta \in [1,3]$. In this framework, p-values represent a statistic testing two crucial aspects of the modelling: sample independence and model compatibility with data. The statistic consists of the D'Agostino K^2 test together with Spearman's rank correlation of residuals, which evaluates compatibility and independence, respectively (Leitão et al., 2016)

Finally, we compare each of the four models individually against the linear alternative (with fixed $\beta = 1$), to test the nonlinearity plausibility. With the fits of all types of crime and countries, we measure the Bayesian information criterion (BIC), defined as

$$\text{BIC} = -2\ln \mathcal{L} + k \ln n,$$

where k is the number of free parameters in the model and lower BIC values indicate better data description. The BIC value of each fit enables us to compare the models' ability to explain data.

Appendices
Appendix 1: Results from the probabilistic scaling analysis
To test the plausibility of a nonlinear scaling, we compare each model against the linear alternative (i.e., $\beta = 1$) using the difference ΔBIC between the fits for each data set. We follow Leitão et al. (2016) and define three outcomes from this comparison. First, if ΔBIC < 0, we say that the model is linear ($\rightarrow$), since we can consider that the linear model explains the data better. Second, if $0 < \Delta$BIC < 6, we consider the analysis of $\beta \neq 1$ inconclusive because we do not have enough evidence for the nonlinearity. Finally, if ΔBIC > 6, we have evidence in favor of the nonlinear scaling, which can be superlinear ($\nearrow$) or sublinear

($\searrow$). We also use ΔBIC to determine the model P(Y|N) that describes the data better. In Tables 2 and Table 3, we summarize the results in that a dark gray cell value indicates the best model based on ΔBIC, a light gray value indicates the best model given a P(Y|N) model, and $*$ indicates that the model is plausible (p-value > 0.05).

Appendix 2: Data sources
See Table 4.

Table 4 Data sources for each country

Country	URL
Belgium	https://bit.ly/3v5l8am
Canada	https://bit.ly/3f42Jq7
Colombia	https://bit.ly/3u7lFZb
Denmark	https://bit.ly/3bETREY
France	https://bit.ly/3u8727Q
Italy	https://bit.ly/2S9iAL0
Mexico	https://bit.ly/3v5JFgC https://bit.ly/3lkch3w
Portugal	https://bit.ly/3hHbMz0
South Africa	https://bit.ly/3u7m6CN
Spain	https://bit.ly/3wqOWjb
United Kingdom	https://bit.ly/3ytMOJ0
United States	https://bit.ly/3fx2iTZ

Declarations

Author details
[1]University of Exeter College of Engineering Mathematics and Physical Sciences, Exeter, UK. [2]GESIS-Leibniz Institute for the Social Sciences, Cologne, Germany.

References

Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K., & Mendes, R. S. (2013a). Distance to the scaling law: A useful approach for unveiling relationships between crime and urban metrics. *PLoS ONE, 8*(8), e69580.

Alves, L. G. A., Ribeiro, H. V., & Mendes, R. S. (2013b). Scaling laws in the dynamics of crime growth rate. *Physica A: Statistical Mechanics and Its Applications, 392*(11), 2672–2679.

Andresen, M. A. (2006). Crime measures and the spatial analysis of criminal activity. *The British Journal of Criminology, 46*(2), 258–285.

Andresen, M. A. (2011). The ambient population and crime analysis. *The Professional Geographer, 63*(2), 193–212.

Andresen, M. A., & Linning, S. J. (2012). The (in)appropriateness of aggregating across crime types. *Applied Geography, 35*(1–2), 275–282.

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2014). Constructing cities, deconstructing scaling laws. *Journal of The Royal Society Interface, 12*(102), 20140745.

Balbi, A., & Guerry, A.-M. (1829). *Statistique comparée de l'état de l'instruction et du nombre des crimes dans les divers arrondissements des Académies et des Cours Royales de France.* Paris: Everat.

Bettencourt, L. M. A. (2013). The origins of scaling in cities. *Science, 340*(6139), 1438–1441.

Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences, 104*(17), 7301–7306.

Bettencourt, L. M. A., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS ONE, 5*(11), e13541.

Bettencourt, L. M. A., Lobo, J., & Youn, H. (2013). The hypothesis of urban scaling: Formalization, implications and challenges. Technical Report arXiv:1301.5919 [physics.soc-ph], arXiv.

Blau, P. (1977). *Inequality and heterogeneity: A primitive theory of social structure* (1st ed.). Free Press.

Boggs, S. L. (1965). Urban crime patterns. *American Sociological Review, 30*(6), 899.

Boivin, R. (2013). On the use of crime rates. *Canadian Journal of Criminology and Criminal Justice, 55*(2), 263–277.

Bursik, R. J., & Webb, J. (1982). Community change and patterns of delinquency. *American Journal of Sociology, 88*(1), 24–42.

Chamlin, M. B., & Cochran, J. K. (2004). An excursus on the population size–crime relationship. *Western Criminology Review, 5*(2), 119–130.

Chang, Y. S., Kim, H. E., & Jeon, S. (2019). Do larger cities experience lower crime rates? A scaling analysis of 758 cities in the U.S. *Sustainability, 11*(11), 3111.

Clarke, R. V. (1984). Opportunity-based crime rates: The difficulties of further refinement. *The British Journal of Criminology, 24*(1), 74–83.

Cohen, L. E., Kaufman, R. L., & Gottfredson, M. R. (1985). Risk-based crime statistics: A forecasting comparison for burglary and auto theft. *Journal of Criminal Justice, 13*(5), 445–457.

Cottineau, C., Hatna, E., Arcaute, E., & Batty, M. (2017). Diverse cities or the systematic paradox of urban scaling laws. *Computers, Environment and Urban Systems, 63*, 80–94.

Durkheim, É. (1895). *Les règles de la méthode sociologique.* Paris: Flix Alcan.

Fischer, C. S. (1975). Toward a subcultural theory of urbanism. *American Journal of Sociology, 80*(6), 1319–1341.

Fischer, C. S. (1995). The subcultural theory of urbanism: A twentieth-year assessment. *American Journal of Sociology, 101*(3), 543–577.

Freudenburg, W. R. (1986). The density of acquaintanceship: An overlooked variable in community research? *American Journal of Sociology, 92*(1), 27–63.

Friendly, M. (2007). A.-M. Guerry's Moral statistics of France : Challenges for multivariable spatial analysis. *Statistical Science 22*(3), 368–399.

Gibbs, J. P., & Erickson, M. L. (1976). Crime rates of American cities in an ecological context. *American Journal of Sociology, 82*(3), 605–620.

Gomez-Lievano, A., Patterson-Lomba, O., & Hausmann, R. (2016). Explaining the prevalence, scaling and variance of urban phenomena. *Nature Human Behaviour, 1*(1), 0012.

Gomez-Lievano, A., Youn, H., & Bettencourt, L. M. A. (2012). The Statistics of Urban Scaling and Their Connection to Zipf's Law. *PLoS ONE, 7*(7), e40393.

Groff, E. R. (2015). Informal Social Control and Crime Events. *Journal of Contemporary Criminal Justice, 31*(1), 90–106.

Hall, J. (2016). Flint is the most dangerous city in America—But it has nothing to do with the water crisis. *The Independent.*

Hanley, Q. S., Khatun, S., Yosef, A., & Dyer, R.-M. (2014). Fluctuation scaling, Taylor's law, and crime. *PLoS ONE, 9*(10), e109004.

Hanley, Q. S., Lewis, D., & Ribeiro, H. V. (2016). Rural to urban population density scaling of crime and property transactions in English and Welsh Parliamentary Constituencies. *PLOS ONE, 11*(2), e0149546.

Harrendorf, S. (2018). Prospects, problems, and pitfalls in comparative analyses of criminal justice data. *Crime and Justice, 47*(1), 159–207.

Harrendorf, S., Heiskanen, M., & Malby, S. (Eds.). (2010). *International statistics on crime and justice*, Volume 64. Helsinki: European Institute for Crime Prevention and Control, affiliated with the United Nations.

Harries, K. D. (1981). Crime rates and environmental denominators. *Journal of Environmental Systems, 11*(1), 3–15.

Hipp, J. R., & Steenbeek, W. (2016). Types of crime and types of mechanisms. *Crime & Delinquency, 62*(9), 1203–1234.

Leitão, J. C., Miotto, J. M., Gerlach, M., & Altmann, E. G. (2016). Is this scaling nonlinear? *Royal Society Open Science, 3*(7), 150649.

Louf, R., & Barthelemy, M. (2014). Scaling: Lost in the smog. *Environment and Planning B: Planning and Design, 41*(5), 767–769.

Malleson, N., & Andresen, M. A. (2016). Exploring the impact of ambient population measures on London crime hotspots. *Journal of Criminal Justice, 46*, 52–63.

Mayhew, B. H., & Levinger, R. L. (1976). Size and the Density of Interaction in Human Aggregates. *American Journal of Sociology, 82*(1), 86–110.

Miethe, T., McCorkle, R., & Listwan, S. J. (2005). *Crime profiles: The anatomy of dangerous persons, places, and situations* (3rd edn.). Oxford University Press.

Miethe, T. D., Hughes, M., & McDowall, D. (1991). Social change and crime rates: An evaluation of alternative theoretical approaches. *Social Forces, 70*(1), 165–185.

Oliveira, M., Barbosa-Filho, H., Yehle, T., White, S., & Menezes, R. (2015). From criminal spheres of familiarity to crime networks. In *Complex Networks VI*, pp. 219–230. Springer.

Oliveira, M., Bastos-Filho, C., & Menezes, R. (2017). The scaling of crime concentration in cities. *PLOS ONE, 12*(8), e0183110.

Oliveira, M., & Menezes, R. (2019). Spatial concentration and temporal regularities in crime. arXiv preprint arXiv:1901.03589.

Oliveira, M., Ribeiro, E., Bastos-Filho, C., & Menezes, R. (2018). Spatio-temporal variations in the urban rhythm: the travelling waves of crime. *EPJ Data Science, 7*(1), 29.

Pacheco, D. F., Oliveira, M., & Menezes, R. (2017). Using social media to assess neighborhood social disorganization: A case study in the United Kingdom. In *Proceedings of the thirtieth international Florida artificial intelligence research society conference, FLAIRS 2017, Marco Island, Florida, USA, May 22–24, 2017.*, pp. 341–346.

Park, H., & Katz, J. (2016). Murder rates rose in a quarter of the nation's 100 largest cities. *The New York Times.*

Rotolo, T., & Tittle, C. R. (2006). Population size, change, and crime in U.S. cities. *Journal of Quantitative Criminology, 22*(4), 341–367.

Sampson, R. J. (1986). Crime in cities: The effects of formal and informal social control. *Crime and Justice, 8*, 271–311.

Sampson, R. J. (1988). Local friendship ties and community attachment in mass society: A multilevel systemic model. *American Sociological Review, 53*(5), 766.

Siegel, L. J. (2011). *Criminology* (11th ed.). Wadsworth Publishing.

Stults, B. J., & Hasbrouck, M. (2015). The effect of commuting on city-level crime rates. *Journal of Quantitative Criminology, 31*(2), 331–350.

Takala, J.-P., & Aromaa, K. (2008). Victimology. In *Encyclopedia of violence, peace, & conflict*, pp. 2272–2288. Elsevier.

Thurner, S., Hanel, R., & Klimek, P. (2018). *Introduction to the theory of complex systems.* Oxford University Press.

White, S., Yehle, T., Serrano, H., Oliveira, M., & Menezes, R. (2014). The spatial structure of crime in urban environments. In *Social informatics*, pp. 102–111. Springer.

Wirth, L. (1938). Urbanism as a Way of Life. *American Journal of Sociology, 44*(1), 1–24.

Yang, V. C., Papachristos, A. V., & Abrams, D. M. (2019). Modeling the origin of urban-output scaling laws. *Physical Review E, 100*(3), 32306.