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,*]

[1]*Computer Science, University of Exeter, Exeter, United Kingdom*

[2]*GESIS–Leibniz Institute for the Social Sciences, Cologne, Germany*

Crime rates per capita are used virtually everywhere to rank and compare cities. However, they rely on a strong linear assumption that crime increases at the same pace as the number of people in a region. Here we show 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 twelve countries and assess the impact of per capita measurements on crime analyses, depending on offense type. 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 in such a way that they disagree about half of the cities in the top ten most dangerous cities. We advise caution when using crime rates per capita to rank cities and recommend evaluating the linear plausibility before analyzing crime rates.

Keywords: crime rate, population size, city, urban scaling, ranking, complex systems

Introduction

In criminology, it is a generally accepted fact 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 removes the *linear* effect of population size from crime numbers, and it 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 neglects 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

* moliveira@tuta.io

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 way to compare crime in cities.

However, the widespread adoption of crime rates is arguably due more to tradition rather than its ability to remove the effects of population size (Boivin 2013). 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 population effects, which introduces an artifactual bias into rankings and analyses.

Despite this inadequacy, we only have a limited understanding of the impact of nonlinearity on crime rates. The literature has mostly paid attention to estimating the relationship between crime and population size, focusing on either specific countries or crime types. The lack of comprehensive systematic studies has limited our knowledge regarding the impact of the linear assumption on 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 twelve 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 then we 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 sub-cultural (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 each other. These theories, however, differ in the type 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 increases the 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 which promotes different social associations such as victimization. At the same time that conflictive association increases, other integrative ones 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 population: size and stability. Larger population size leads to higher population density and heterogeneity—not only individuals have more opportunities for social contacts, but they are often surrounded by strangers (Wirth 1938). This situation makes social integration more difficult and promotes a higher anonymity, which encourages criminal impulses and harms 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). This viewpoint understands 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, increasing misbehavior and crime (Miethe et al. 1991; Sampson 1988).

Both social control and structural perspectives solely focus on individuals' interactions without considering individuals' 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 *sub-cultural* perspective advocates that population concentration brings together individuals with shared interests, which produces private social networks built around these interests, promoting a 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 sufficient people with specific shared interests which enable 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 sub-cultural—expect that more people in an area lead 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 rate per capita

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

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

which are often used together with a multiplier that contextualizes the quantity (e.g., crime per $100,000$ inhabitants) (Boivin 2013). However, though crime rates are popularly used, they present at least two inadequacies. First, the way 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 to use ambient population (Andresen 2006, 2011) and account for the number of targets, which depends on the type of crime (Boggs 1965; Cohen et al. 1985).

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. Because of the linearity 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 of population. Not all theoretical perspectives, however, agree with such a 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 2013; Bettencourt et al. 2007, 2010; Gomez-Lievano et al. 2016). Bettencourt et al. (2007) have shown that a city's population size, denoted by N, is a strong predictor of its urban indicators, denoted by Y, exhibiting a relationship of the form:

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

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 of their idiosyncrasies (Oliveira and Menezes 2019). To understand this scaling and 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 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), there exists an economy of scale. 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 linearity assumption of per capita ratios, we deal with quantities that can misrepresent an urban indicator. To show that, we use Eq. (3) to define the per capita rate C of an urban indicator as the following:

$$C = \frac{Y}{N} \sim N^{\beta-1}, \tag{4}$$

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) (Alves et al. 2013; Bettencourt et al. 2010). This misrepresentation occurs because population still has an effect on rates. By definition, we expect that per capita rates are higher in bigger cities when $\beta > 1$, whereas when $\beta < 1$, we expect bigger cities having lower rates. In nonlinear situations, when we compare cities via rates, we introduce an artifactual bias in analyses and rankings of urban

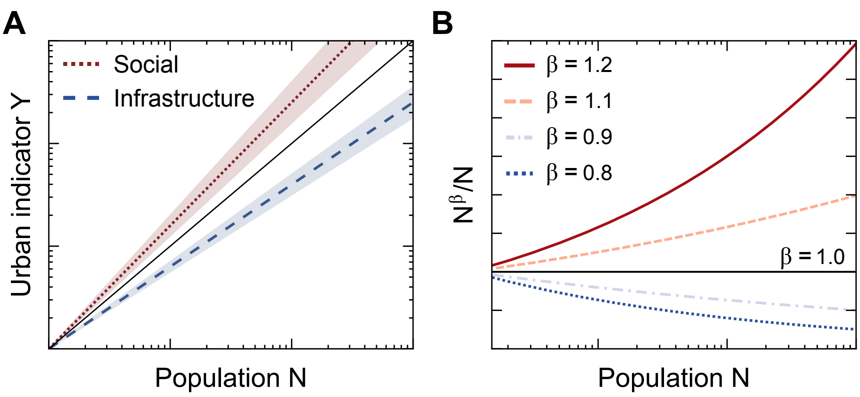

Fig. 1. **The urban scaling laws and rates per capita.** The way urban indicators increase with population size depends on the class of the indicator. (**A**) Social aspects, such as crime and total wages, increase superlinearly with population size, whereas infrastructural indicators (e.g., road length) increase sublinearly. (**B**) In nonlinear scenarios, rates per capita still depend on population size.

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 a 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. 2013; Gomez-Lievano et al. 2012). Previous works have also shown that different kinds of crime in the U.K. and in U.S. present nonlinear scaling relationships (Chang et al. 2019; Hanley et al. 2016). Remarkably, the existence of these scaling laws of crime suggests fundamental urban processes that relates 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:

$$Y(\kappa N) = g(\kappa)Y(N), \tag{5}$$

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 of crime in such a way that $g(\kappa) = \kappa^{\beta}$. Second, Eq. (5) 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). In general, 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 a regularity is independent of crime type.

However, different types of crime are connected to social mechanisms differently (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 twelve 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 the following:

$$\text{crime} \sim f(\text{population}). \tag{6}$$

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 linearity assumption).

Results

We use data from twelve countries to investigate the relationship between population size and crime at the city level. We examine annual data from Belgium, Canada, Colombia, Denmark, France, Italy, Portugal, South Africa, Spain, the United Kingdom, and the United States (see Table I). In our research, we are not interested in comparing countries' absolute numbers of crime. We understand that international comparisons of crime have several problems because of differences in crime definitions, police and court practices, reporting rates, and others (Takala and Aromaa 2008). In this work, we want to investigate how crime increases with population size in each country, focusing on burglary and theft (see Supplementary Information for data sources). We analyze data of both types of crime in all considered countries, except Mexico, Portugal, and Spain, where we only have data for one kind of offense.

TABLE I. Burglary and theft annual statistics in twelve countries: number of data points n, sample mean $\bar{y}$, sample standard deviation S, and maximum value y_{max}.

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	37150	293.90	791.13	7782
Colombia	513	182.04	1514.68	36306	40.08	228.06	4856
Denmark	98	1157.67	3851.29	38011	330.71	330.60	2157
France	100	8311.12	12400.34	108846	2389.94	2515.24	12511
Italy	107	17470.72	30860.27	218052	2217.50	2642.61	18101
Mexico	1659	237.56	959.59	14999	-	-	-
Portugal	279	-	-	-	51.38	86.91	850
South Africa	199	2305.23	8758.52	93793	1190.03	3212.93	28143
Spain	144	7846.72	25111.99	236026	-	-	-
United Kingdom	313	1763.43	1965.61	19766	620.98	685.40	4825
United States	8337	471.82	2345.27	108376	127.33	626.16	19859

The scaling laws of crime in cities

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

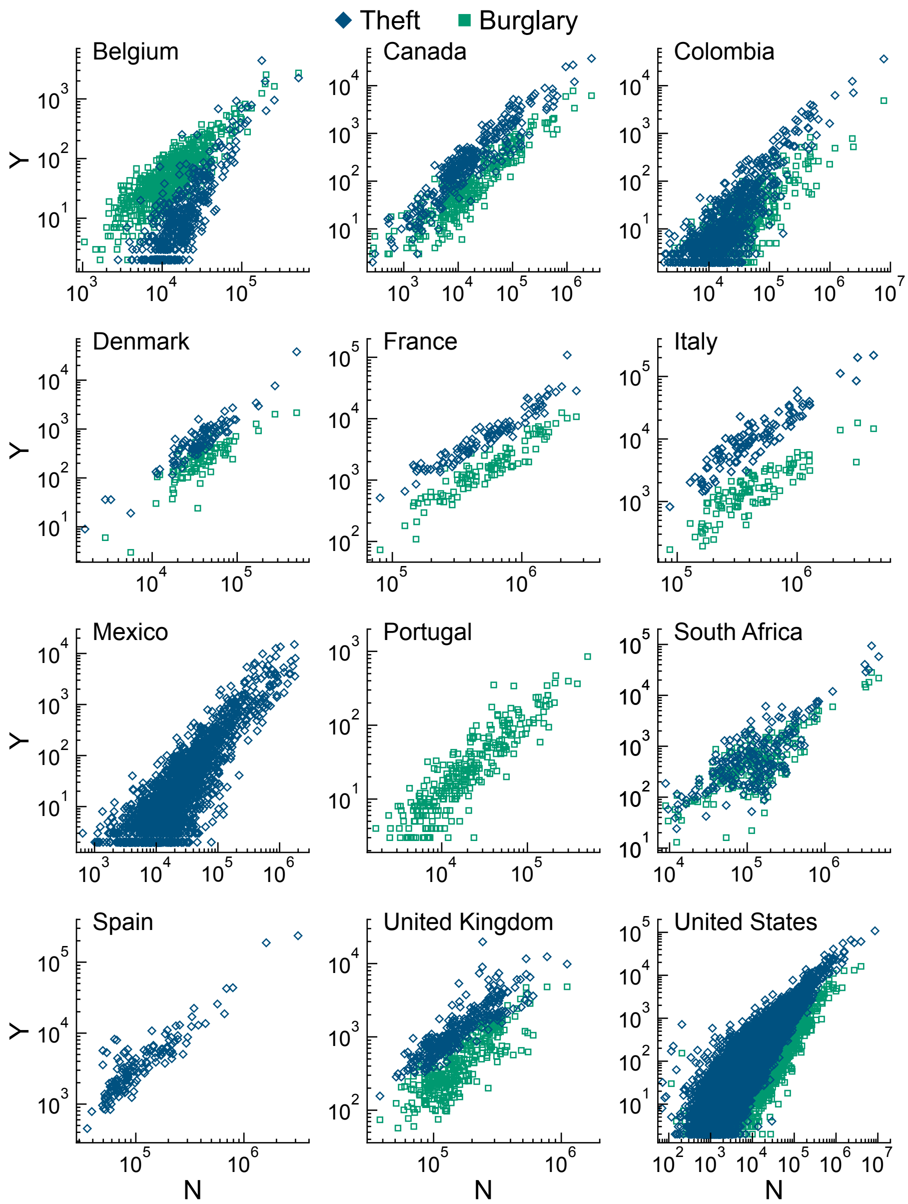

Fig. 2. **The population–crime relationship in twelve countries.** Different criminology theories expect a relationship between population size and crime. They predict, however, divergent population effects, such as linear and superlinear crime growth. Yet, crime rates per capita assume a linear crime growth.

In most of the considered countries, theft increases with population size superlinearly, whereas burglary tends to increase linearly (see Fig. 3). Precisely, in nine out of eleven 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 seven out of ten 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 we have data for different years (see Appendix I).

Our results show 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 (> 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

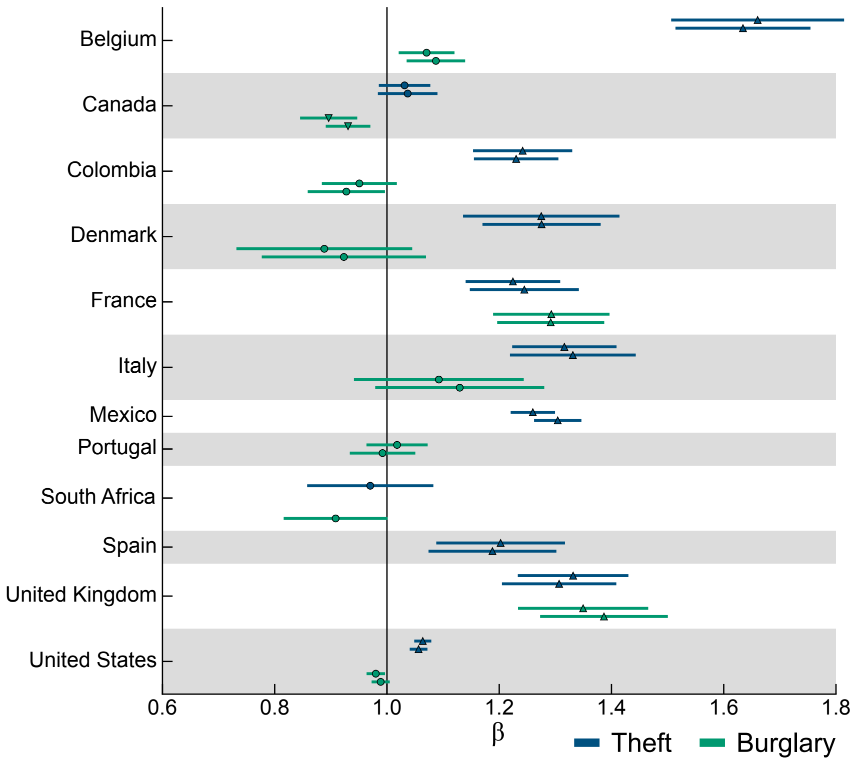

Fig. 3. **The scaling laws of crime.** We find evidence for a nonlinear relationship between crime and population size in more than half of the data sets. In most considered countries, theft exhibits superlinearity, whereas burglary tends to display linearity. In the plot, the lines are the error bars 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.

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 implies that analyses with a linear assumption might create distorted pictures of crime in cities. To understand such distortions, we have to 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 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 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 ones of larger cities. We have to account for this tendency in order to compare crime in cities; otherwise, we introduce bias against

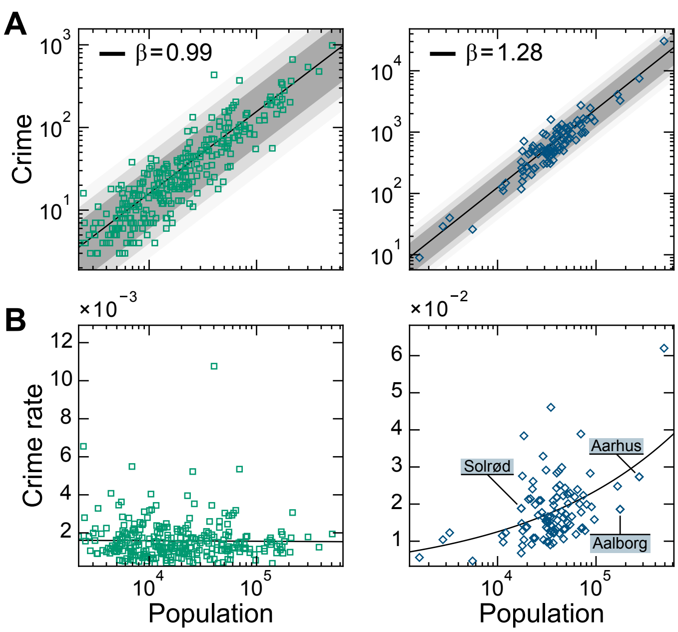

Fig. 4. **Bias in crime rates per capita.** When crime increases nonlinearly with population size, we have an artifactual bias in crime rates. The linearity in Portugal makes rates independent of size (left). In Denmark, however, we expect bigger cities to have higher crime *rates* due to superlinear growth (right). 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.

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 crime 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 tells us how much more or less crime a particular city has in comparison to cities with 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. We denote this kind of analysis as a comparison adjusted for 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 the expected for cities of similar size, while crime in Solrød is above the model expectation (see Fig. 4B). This disagreement arises because of the different population sizes. Since Aalborg is more than ten 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 crime rankings of cities. The municipality of Aarhus, in Denmark, for example, is in the top twelve ranking of 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 fifty-four ranking.

To understand these variations systematically, we compare rankings based on crime rates with rankings that account for population–crime relationship (i.e., adjusted rankings). Our results show 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 then examine the change in the rank of each city. We

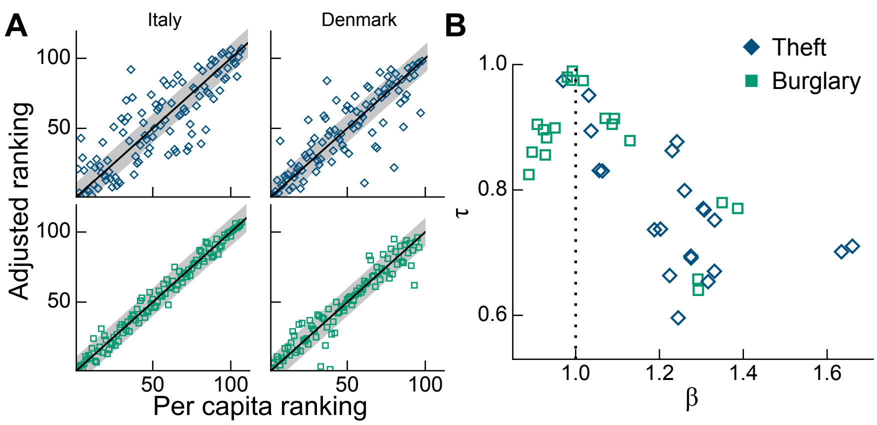

Fig. 5. **The inadequacy of per capita rankings.** Per capita ranking can differ substantially from rankings adjusted for population size, depending on the scaling exponent. In Italy and Denmark, for example, (**A**) theft ranks (top) diverge considerably more than the ranks for burglary (bottom). Data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

find that 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 eleven positions from the adjusted ranking (Fig. 5A). This disagreement means that these rankings disagree about half of the cities in the top ten 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 to 1, the rankings are more similar to each other. For example, in Italy, in contrast to theft, the burglary rate rank of half of the cities only differs from the adjusted ranking in a maximum of two positions (Fig. 5A).

Discussion

Despite being used virtually everywhere, crime rates per capita have a strong assumption that crime increases at the same pace as the number of people in a region. In this work, we investigated how crime grows with population size and how such a widespread assumption of linear growth influences cities' rankings.

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

We found 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. Our results showed that using crime rates to rank cities substantially differs from ranking cities while adjusting for population size.

From academia to news outlets, crime rates per capita are arguably used because they provide us with a familiar measure of criminality (Boivin 2013). Our work implies, however, that they can create a distorted picture of crime in cities. For example, in superlinear scenarios, we expect bigger cities to have higher crime *rates*. In this case, when we use rates to rank cities, we build rankings that big cities are at the top. But, these cities might not experience more crime than what we expect from places of the same size. It is an artifactual bias due to population effects still present in crime rates.

Because of this inadequacy, we advise caution when using crime rates per capita to compare cities. We recommend first evaluating the linear plausibility before analyzing crime rates, and avoiding them when possible. Instead, we suggest comparing z scores computed via the model estimated using the approach discussed in the manuscript (Leitão et al. 2016).

We highlight that crime rates per capita also suffer from the population definition issue—that is, how we define population affects crime rates. In this work, we used the resident population to analyze the population–crime relationship. We understand that crime is not limited to residents (Gibbs and Erickson 1976), and cities attract non-residents (Stults and Hasbrouck 2015). Much literature suggests using ambient population and account for the number of targets (Andresen 2006, 2011; Boggs 1965). However, this data is difficult to collect when dealing with different countries. Future research should investigate the scaling laws using other definitions of population, particularly using social media data (Malleson and Andresen 2016; Pacheco et al. 2017).

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. We need better models—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.

Data and methods

Preprocessing data

We gathered data sets of different types of crime at the city level from 12 countries: Belgium, Canada, Colombia, Denmark, France, Italy, Portugal, South Africa, Spain, United Kingdom, and United States. To examine different types of crime in these countries, we need to have a way to denote each type of crime in each place using a general description. The way we categorize the different types of crime are summarized in the Supplementary Material.

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:

$$\mathrm{E}[Y|N] = \lambda N^{\beta}, \tag{7}$$

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 $\mathrm{P}(Y|N)$ that satisfy the following conditional variance:

$$\mathrm{V}[Y|N] = \gamma \mathrm{E}[Y|N]^{\delta}, \tag{8}$$

where typically $\delta \in [1,2]$. To estimate the scaling laws, we maximize the log-likelihood

$$\mathscr{L} = \ln \mathrm{P}(y_1, \ldots, y_K | n_1, \ldots, n_K) = \sum_{i=1}^{K} \ln \mathrm{P}(y_i | n_i), \tag{9}$$

since we assume y_i as an independent realization from $\mathrm{P}(Y|N)$. 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 $\mathrm{P}(Y|N)$ using Gaussian and log-normal distributions, so we can analyze whether accounting for the size-dependent variance influences the estimation. In the case of the Gaussian, the conditions from Eq. (7) and Eq. (8) are satisfied with

$$\mu_{\mathsf{N}}(x) = \alpha x^{\beta} \quad \text{and} \quad \sigma^2_{\mathsf{N}}(x) = \gamma(\alpha x^{\beta})^{\delta}, \tag{10}$$

whereas in the case of the log-normal distribution,

$$\mu_{\mathsf{LN}}(x) = \ln \alpha + \beta \ln x - \frac{1}{2}\sigma^2_{\mathsf{LN}}(x) \quad \text{and} \quad \sigma^2_{\mathsf{LN}}(x) = \ln\left[1 + \gamma(\alpha x^{\beta})^{\delta-2}\right]. \tag{11}$$

In log-normal case, note that, if $\delta = 2$, 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 that δ is also included in the optimization process. In the case of the Gaussian, we have fixed $\delta = 1$ and free $\delta \in [1,2]$. In the case of the log-normal, we have fixed $\delta = 2$ and free $\delta \in [1,3]$.

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 Criteria (BIC), defined as

$$\mathrm{BIC} = -2 \ln \mathscr{L} + k \ln n, \tag{12}$$

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 ability of the models to explain data.

Declarations

Competing interests. The authors declare that they have no competing interests.

Funding. No outside funding was used to support this work.

Author contributions. All authors read and approved the final manuscript.

Availability of data and materials. The final submission will be accompanied by source code, data, and a live tutorial on the Web where researchers and practitioners can analyze crime using their data.

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

Andresen, M. A. (2006, mar). Crime Measures and the Spatial Analysis of Criminal Activity. *The British Journal of Criminology 46*(2), 258–285.

Andresen, M. A. (2011, mar). The Ambient Population and Crime Analysis. *The Professional Geographer 63*(2), 193–212.

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

Balbi, A. and A.-M. Guerry (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, jun). The origins of scaling in cities. *Science 340*(6139), 1438–1441.

Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kühnert, and G. B. West (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., J. Lobo, D. Strumsky, and G. B. West (2010, nov). Urban Scaling and Its Deviations: Revealing the Structure of Wealth, Innovation and Crime across Cities. *PLoS ONE 5*(11), e13541.

Bettencourt, L. M. A., J. Lobo, and H. Youn (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* (1 ed.). Free Press.

Boggs, S. L. (1965, dec). Urban Crime Patterns. *American Sociological Review 30*(6), 899.

Boivin, R. (2013, apr). On the Use of Crime Rates. *Canadian Journal of Criminology and Criminal Justice 55*(2), 263–277.

Bursik, R. J. and J. Webb (1982, jul). Community Change and Patterns of Delinquency. *American Journal of Sociology 88*(1), 24–42.

Chamlin, M. B. and J. K. Cochran (2004). An Excursus on the Population Size-Crime Relationship. *Western Criminology Review 5*(2), 119–130.

Chang, Y. S., H. E. Kim, and S. Jeon (2019, jun). Do Larger Cities Experience Lower Crime Rates? A Scaling Analysis of 758 Cities in the U.S. *Sustainability 11*(11), 3111.

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

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

Fischer, C. S. (1975, may). Toward a Subcultural Theory of Urbanism. *American Journal of Sociology 80*(6), 1319–1341.

Fischer, C. S. (1995, nov). The Subcultural Theory of Urbanism: A Twentieth-Year Assessment. *American Journal of Sociology 101*(3), 543–577.

Freudenburg, W. R. (1986, jul). The Density of Acquaintanceship: An Overlooked Variable in Community Research? *American Journal of Sociology 92*(1), 27–63.

Friendly, M. (2007, aug). A.-M. Guerry's Moral Statistics of France : Challenges for Multivariable Spatial Analysis. *Statistical Science 22*(3), 368–399.

Gibbs, J. P. and M. L. Erickson (1976, nov). Crime Rates of American Cities in an Ecological Context. *American Journal of Sociology 82*(3), 605–620.

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

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

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

Hall, J. (2016, Jan). Flint is the most dangerous city in america - but it has nothing to do with the water crisis. *The Independent*.

Hanley, Q. S., D. Lewis, and H. V. Ribeiro (2016, feb). Rural to Urban Population Density Scaling of Crime and Property Transactions in English and Welsh Parliamentary Constituencies. *PLOS ONE 11*(2), e0149546.

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

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

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

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

Miethe, T., R. McCorkle, and S. J. Listwan (2005). *Crime Profiles: The Anatomy of Dangerous Persons, Places, and Situations* (3 ed.). Oxford University Press.

Miethe, T. D., M. Hughes, and D. McDowall (1991, sep). Social Change and Crime Rates: An Evaluation of Alternative Theoretical Approaches. *Social Forces 70*(1), 165–185.

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

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

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

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

Pacheco, D. F., M. Oliveira, and R. Menezes (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. and J. Katz (2016, Sep). Murder rates rose in a quarter of the nations 100 largest cities. *The New York Times*.

Rotolo, T. and C. R. Tittle (2006, oct). Population Size, Change, and Crime in U.S. Cities. *Journal of Quantitative Criminology 22*(4), 341–367.

Sampson, R. J. (1986, jan). Crime in Cities: The Effects of Formal and Informal Social Control. *Crime and Justice 8*, 271–311.

Sampson, R. J. (1988, oct). Local Friendship Ties and Community Attachment in Mass Society: A Multilevel Systemic Model. *American Sociological Review 53*(5), 766.

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

Stults, B. J. and M. Hasbrouck (2015, jun). The Effect of Commuting on City-Level Crime Rates. *Journal of Quantitative Criminology 31*(2), 331–350.

Takala, J.-P. and K. Aromaa (2008). Victimology. In *Encyclopedia of Violence, Peace, & Conflict*, pp. 2272–2288. Elsevier.

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

White, S., T. Yehle, H. Serrano, M. Oliveira, and R. Menezes (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.

Appendices

Appendix I: 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 $\Delta\text{BIC} < 0$, we say that the model is linear ($\rightarrow$), since we can consider that the linear model explains the data better. Second, if $0 < \Delta\text{BIC} < 6$, we consider the analysis of $\beta \neq 1$ inconclusive because we do not have enough evidence for the nonlinearity. Finally, if $\Delta\text{BIC} > 6$, 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 $\text{P}(Y|N)$ that describes the data better. In Table II and Table III, we summarize the results in that we a dark gray cell indicates the best model based on ΔBIC, a light gray cell indicates the best model given a $\text{P}(Y|N)$ model, and $*$ indicates that the model is plausible (> 0.05).

TABLE II. 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) $\nearrow$	1.64 (0.12) $\nearrow$	2.11 (0.27) $\nearrow$	1.67 (0.17) $\nearrow$
Belgium (2016)	1.66 (0.15) $\nearrow$	1.66 (0.14) $\nearrow$	2.10 (0.18) $\nearrow$	1.75 (0.19) $\nearrow$
Canada (2015)	1.09 (0.06) $\nearrow$	1.04 (0.05) $\rightarrow$	1.07 (0.11) $\rightarrow$	1.04 (0.06) $\rightarrow$
Canada (2016)	1.03 (0.04) $\rightarrow$	1.04 (0.05) $\rightarrow$	1.06 (0.34) $\circ$	1.03 (0.05) $\rightarrow$
Colombia (2013)	1.25 (0.07) $\nearrow$	1.23 (0.07) $\nearrow^*$	1.89 (0.09) $\nearrow$	1.31 (0.08) $\nearrow$
Colombia (2014)	1.26 (0.07) $\nearrow$	1.24 (0.09) $\nearrow^*$	1.89 (0.09) $\nearrow$	1.36 (0.08) $\nearrow$
Denmark (2015)	1.28 (0.10) $\nearrow^*$	1.27 (0.13) $\nearrow^*$	1.45 (0.33) $\nearrow$	1.27 (0.14) $\nearrow$
Denmark (2016)	1.27 (0.14) $\nearrow^*$	1.28 (0.18) $\nearrow^*$	1.58 (0.37) $\nearrow$	1.28 (0.18) $\nearrow$
France (2013)	1.24 (0.09) $\nearrow$	1.23 (0.07) $\nearrow^*$	1.59 (0.44) $\nearrow$	1.30 (0.12) $\nearrow$
France (2014)	1.24 (0.10) $\nearrow$	1.22 (0.08) $\nearrow$	1.70 (0.57) $\nearrow$	1.34 (0.18) $\nearrow$
Italy (2014)	1.33 (0.11) $\nearrow^*$	1.31 (0.10) $\nearrow^*$	1.37 (0.15) $\nearrow$	1.31 (0.09) $\nearrow$
Italy (2015)	1.32 (0.09) $\nearrow^*$	1.29 (0.11) $\nearrow^*$	1.35 (0.14) $\nearrow$	1.29 (0.10) $\nearrow$
Mexico (2015)	1.30 (0.04) $\nearrow$	1.31 (0.04) $\nearrow$	1.98 (0.02) $\nearrow$	1.32 (0.04) $\nearrow$
Mexico (2016)	1.26 (0.04) $\nearrow$	1.26 (0.04) $\nearrow$	1.98 (0.01) $\nearrow$	1.30 (0.05) $\nearrow$
South Africa (2016)	0.97 (0.11) $\rightarrow^*$	0.99 (0.10) $\rightarrow^*$	1.33 (0.20) $\nearrow$	1.02 (0.11) $\rightarrow$
Spain (2015)	1.18 (0.11) $\nearrow$	1.19 (0.11) $\nearrow$	1.27 (0.19) $\nearrow$	1.22 (0.12) $\nearrow$
Spain (2016)	1.20 (0.11) $\nearrow$	1.20 (0.11) $\nearrow$	1.31 (0.20) $\nearrow$	1.24 (0.13) $\nearrow$
United Kingdom (2015)	1.24 (0.07) $\nearrow$	1.31 (0.10) $\nearrow$	1.45 (0.30) $\nearrow$	1.55 (0.32) $\nearrow$
United Kingdom (2016)	1.26 (0.09) $\nearrow$	1.33 (0.10) $\nearrow$	1.50 (0.37) $\nearrow$	1.59 (0.35) $\nearrow$
United States (2014)	1.12 (0.01) $\nearrow$	1.06 (0.01) $\nearrow$	1.07 (0.06) $\nearrow$	1.04 (0.04) $\nearrow$
United States (2015)	1.13 (0.01) $\nearrow$	1.06 (0.01) $\nearrow$	1.08 (0.07) $\nearrow$	1.05 (0.04) $\nearrow$

TABLE III. 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) $\nearrow$	1.09 (0.05) $\circ$	1.21 (0.11) $\nearrow$	1.09 (0.05) $\circ$
Belgium (2016)	1.08 (0.06) $\circ$	1.07 (0.05) $\circ$	1.18 (0.10) $\nearrow$	1.08 (0.05) $\circ$
Canada (2015)	0.93 (0.05) $\circ$	0.93 (0.04) $\searrow^*$	1.04 (0.10) $\rightarrow$	0.95 (0.06) $\circ$
Canada (2016)	0.91 (0.04) $\searrow^*$	0.90 (0.05) $\searrow^*$	1.00 (0.10) $\rightarrow$	0.90 (0.04) $\searrow$
Colombia (2013)	0.90 (0.07) $\circ^*$	0.93 (0.07) $\rightarrow$	1.18 (0.44) $\rightarrow$	0.96 (0.07) $\rightarrow$
Colombia (2014)	0.94 (0.07) $\rightarrow^*$	0.95 (0.06) $\rightarrow$	1.16 (0.51) $\rightarrow$	0.99 (0.07) $\rightarrow$
Denmark (2015)	1.11 (0.26) $\rightarrow$	0.91 (0.14) $\rightarrow$	0.92 (0.14) $\rightarrow^*$	0.93 (0.13) $\rightarrow^*$
Denmark (2016)	1.15 (0.24) $\rightarrow$	0.89 (0.15) $\rightarrow$	0.90 (0.13) $\rightarrow$	0.92 (0.17) $\rightarrow$
France (2013)	1.29 (0.09) $\nearrow^*$	1.27 (0.09) $\nearrow^*$	1.31 (0.11) $\nearrow^*$	1.27 (0.09) $\nearrow$
France (2014)	1.29 (0.10) $\nearrow^*$	1.27 (0.10) $\nearrow^*$	1.34 (0.10) $\nearrow$	1.27 (0.09) $\nearrow$
Italy (2014)	1.13 (0.15) $\rightarrow^*$	1.11 (0.16) $\rightarrow^*$	1.09 (0.17) $\rightarrow$	1.09 (0.12) $\rightarrow^*$
Italy (2015)	1.09 (0.15) $\rightarrow^*$	1.07 (0.13) $\rightarrow^*$	1.06 (0.15) $\rightarrow^*$	1.05 (0.12) $\rightarrow^*$
Portugal (2015)	0.99 (0.06) $\rightarrow^*$	0.98 (0.05) $\rightarrow^*$	1.13 (0.13) $\rightarrow$	0.99 (0.10) $\rightarrow$
Portugal (2016)	1.02 (0.05) $\rightarrow$	1.01 (0.06) $\rightarrow$	1.11 (0.09) $\circ$	1.05 (0.10) $\rightarrow$
South Africa (2016)	0.91 (0.09) $\rightarrow$	0.91 (0.08) $\rightarrow$	1.07 (0.09) $\circ$	0.97 (0.12) $\rightarrow$
United Kingdom (2015)	1.39 (0.11) $\nearrow^*$	1.42 (0.10) $\nearrow^*$	1.47 (0.13) $\nearrow$	1.40 (0.10) $\nearrow$
United Kingdom (2016)	1.35 (0.11) $\nearrow^*$	1.36 (0.10) $\nearrow^*$	1.46 (0.14) $\nearrow$	1.37 (0.11) $\nearrow$
United States (2014)	0.99 (0.01) $\rightarrow$	0.99 (0.01) $\rightarrow$	1.19 (0.11) $\nearrow$	1.07 (0.05) $\nearrow$
United States (2015)	0.98 (0.01) $\rightarrow$	0.98 (0.01) $\circ$	1.17 (0.08) $\nearrow$	1.07 (0.06) $\nearrow$