PAPERS IN APPLIED GEOGRAPHY
https://doi.org/10.1080/23754931.2017.1373257

RESEARCH ARTICLE

Some Reasons for Using Zipf's Law in the Analysis of Urban Crime: The Case of Mexico

Carlos J. Vilalta[a] and Gustavo Fondevila[b]

[a]Center for Research in Geography and Geomatics (CentroGeo), Lomas de Padierna, CP, Mexico City; [b]Center for Economic Research and Education (CIDE), Lomas de Santa Fe, CP, Mexico City

ABSTRACT
The high concentration of crime in a handful of cities is clear. What is not clear, however, is why crime levels are high in particular places. Using crime victimization data from thirty-two Mexican cities, I test one proposition and develop another. First, I show that Zipf's law fits to a certain degree the distribution of crime victimization and fear of crime. Second, I argue that these patterns are due to opportunities to commit crimes that are created not only by the spatial-temporal convergence of potential victims and offenders, but to the principle of least effort, which helps explain urban development processes. In this regard, more and better conditions of production in a city mean not only more jobs and wealth, but also more crime victims. This suggests that urban crime is not only predictable, to some degree, but also a problem associated with the forces of urbanization.

KEYWORDS
Cities; crime; fear of crime; Mexico

Mexican cities differ in many respects, but their similarities could be summed up under one heading: a high prevalence of crime. This is why, among other things, Mexican civil society and the federal government have engaged in heated debates about the functioning of the national crime prevention program, with particular attention being paid to the geographical definition of priority areas (*i.e., demarcaciones prioritarias*) for crime prevention. The premise of the program is that social crime prevention works better in areas with a high risk of crime (*i.e.*, municipalities with high levels of homicide rates). The debate about the geographical definition of priority areas starts with the fact that the methodology for the selection of these municipalities is unclear (Mexico Evalua 2015). The two main disagreements concern whether the selected areas are actually the most dangerous areas, and whether homicide rates should be the only indicator of the need for social crime prevention policy action. What this debate shows, among other things, is that Mexico has a crime data problem. Although Mexico has some administrative records of crime at the municipal level, there currently exist very few crime data using units of analysis greater than (*i.e.*, city or metro area) or smaller than (*i.e.*, neighborhood) the level of the municipality, which has led to some of this definitional debate.

Although the debate has largely focused on addressing problems in the existing data, two additional problems have not yet been debated. One is that the only measure of the impact of the crime prevention program is the rate of victimization or crime prevalence at the national level. Therefore, this measure does not say much, if anything, about the distribution of crime in particular locations within the country. One city with 1 million victims can be confused with ten cities with 100,000 victims, as long as they have the same crime prevalence. These areas are entirely different from a public policy perspective, however. In addition, the current budget formula in federal crime prevention policy prioritizes

CONTACT Carlos J. Vilalta cjvilalta@outlook.com Center for Research in Geography and Geomatics (CentroGeo), Lomas de Padierna, CP 14240, Mexico City; Gustavo Fondevila Center for Economic Research and Education (CIDE), Lomas de Santa Fe, CP 01210, Mexico City.

places based on homicide rates, not actual numbers of crimes committed or the number of victims in general (*i.e.*, the incidence and concentration for all crimes). What is missing is an empirical framework that could help arrange facts in some coherent way.

Unfortunately, within the literature on Latin America, there is little research about the geography of crime and the patterns that emerge among systems of cities (Gaviria and Pagés 2002; Vilalta 2012), although most crimes actually occur in cities. Yet in Mexico in 2014, for example, the largest thirty-two cities in the country accounted for 58 percent (or 13.3 million) of all crime victims nationwide and 66.7 percent (or 23 million) of all crimes. This article makes the case that the city is well suited to the type of analysis needed to address previous debates and that Zipf's law indicates a general property of urban crime quantities. Zipf's law is a probabilistic discovery that states that the frequencies of many social and natural events are inversely proportional to their rank number. For instance, in urban demographics, it indicates that for most countries the size distribution of cities ordered by population fits a power law: The number of cities with populations greater than S is proportional to $1/S$ (Gabaix 1999). With the inevitable trend toward more urbanization, Mexico faces the urgent challenge of developing an accurate assessment of urban victimization.

This article presents a first effort to expand the analysis of urban crime, with the city as a unit of analysis. The article uses recent official survey data from thirty-two Mexican cities to analyze the distribution of victimization and fear of crime. Focusing on interurban variations in victimization rates and individuals reporting perceived likelihood of victimization, this article tests Zipf's law and shows it to be a useful diagnostic tool for detecting crime distributions. The results of this analysis suggest that crime in urban Mexico is mainly a problem of a few cities with millions of crime victims.

This article does not resolve the measurability problems associated with the previous debates, nor it is intended to give an answer to the least effort principle or Zipf's law. Rather, it is intended to help situate it by facilitating the observation of empirical regularities and spatial distributions of crime for a more accurate interpretation of the crime problem. The main finding is that the distributions of the cross-section of crime victimization of Mexican cities in 2011 and 2015 are strongly skewed to the right. This finding seems to suggest an underlying distribution of locational advantages for criminal behavior. It also suggests that Zipf's law has a role to play in the analysis of urban crime distribution.

In sum, Zipf's law is much more than a mathematical oddity. It is able to predict an even falling off of the victimization problem across Mexican cities according to their statistical rank. The results are even more enlightening, as they show Zipf's law to perform better with larger samples rather than smaller samples, and reveal its potential to predict victimization distributions as long as the sample size is sufficient. These results are then discussed in terms of future research in urban crime.

The rank size rule, Zipf's law, and the principle of least effort

In this section, I first introduce the basic formulation of the rank-size rule and Zipf's law. I then provide a summary of the theoretical logic behind the law, the principle of least effort (Zipf 1949).

The rank-size rule is a power law.[1] It states that aggregated data of human behavior follow a distribution of values such that the highest to the lowest values follow one another in harmonic progression (*i.e.*, 100, 50, 33.3, 25, 20, …) when items are ranked in the order of magnitude. It takes the following form: Let R_i, $i = 1 \ldots n$ be the number of crime victims in a system of cities, and $R(_i)$ be its statistical rank, so that

$$\log R_i \approx b_0 + b_1 \log i, \quad i = 1, \ldots, n, \tag{1}$$

where $R(_1) \geq R(_2) \geq \cdots \geq R(_n)$, $b_0 > 0$, and $b_1 < 0$.

When the coefficient of $b_1 \approx -1.0$, it is referred to as Zipf's law. When plotted, this coefficient creates a straight downward sloping line of about –45°. Zipf's law follows the rank-size rule by which the product of the rank and the size of a city is a constant. This law predicts that a large number of crime victims will concentrate in a few large cities and will decrease in a rapid linear fashion as a function of statistical rank order. It should be noted that the law can be stated in both weaker and stronger forms

(Chen 1978); that is, there are some deviations to the −1 slope. A weaker form would be one that deviates passably from the −1 slope. A stronger form would be one that trivially deviates from the −1 slope.

There are two parts to Zipf's law (Gabaix 1999): First, it is a power law and second, it is a power law with a slope near −1.0. The first part is not enigmatic, although the second is not clear at all. Power laws (*i.e.*, Zipf, Pareto, Gibrat, Taylor, Benford, *etc.*) exist due to two simple statistical mathematics principles: scale and base invariance. In simple terms, scale invariance means that if a data set (*i.e.*, victimization rates) following Zipf's law is multiplied by a nonzero constant, the resulting distribution will also follow Zipf's law where b_1 is the scaling exponent. Base invariance means that the type of data transformation or weighting does not matter. Power laws can be found in many probability distributions of random variables, particularly those with long or heavy tails (*i.e.*, Student's t distribution or the Cauchy distribution).

Zipf's law can be seen as a random growth process (Gabaix 1999; Rossi-Hansberg and Wright 2007). In a random growth process, large cities arise due to long histories of favorable economic productivity shocks (Hsu 2008); however, because the probability of such events is low, there are not many large cities and therefore, not many cities with many crime victims. One criticism of this literature is that there is not much explanation regarding what drives the size differences of cities, and whether it is simply randomness, either in urban growth processes (Hsu 2008) or something else.

The question concerning Zipf's law, therefore, is really reduced to the question of why the slope equals −1.0. It follows, as stated earlier, that there are deviations to Zipf's law of a slope equal to −1.0. It is important, however, to be clear with the meaning of the slope of Zipf's law. Systems of cities with $b_1 < -1.0$ represent areas with higher than expected concentrations of victims or large victim quantities in larger cities, whereas those with $b_1 > -1.0$ represent areas with lower than expected concentrations of victims in larger cities.[2] The steeper the slope, the higher the concentration of victims in a few cities.

The empirical regularity of the classic rank-size rule was given a theoretical basis by Zipf (Chang 2016). Zipf formulated the principle of least effort as the primary principle governing entire individual and collective behavior. The idea is that "a person in solving his immediate problems will view these against the background of his probable future problems as estimated by himself … [and to] minimize the total work that he must expend in solving both his immediate problems and his probable future problems … [he will] minimize the probable average rate of his work-expenditure" (Zipf 1949, 1). Briefly put, individuals will tend to minimize their efforts or do the least work required for any task. Zipf's law and the idea of the path of least resistance sprang from this principle. Zipf argued, for instance, that highways tend to represent the shortest, quickest, and easiest path between two cities, just like individuals prefer to use one tool for all tasks, or multitasking tools with reconfigurable parts, rather than a variety of tools for different tasks (Zipf 1949). As can be imagined, there is an economic advantage in decreasing the number of tools, and similarly, there is an economic disadvantage in increasing the number of tools.

Although this is a well-formulated theory, assessing its relationship with city size or other aspects of human aggregated behavior, such as crime, is not easy. Why would city populations and victimization quantities comply with Zipf's law? One possibility can be found in the insights of central place theory (Christaller 1966). Central place theory predicts a scaling principle; that is, city systems are monocentric by definition. They are nodal regions containing a main central city and several surrounding cities of different hierarchical orders that are part of the main central city market area (Burger and Meijers 2012). According to this theory, city systems are characterized by a hierarchy of cities following a rank order based on the size of their market areas and the number of economic functions provided. This idea, applied to victimization, means that crime quantities should differ in the degree of scale economies (*i.e.*, the average cost of committing one crime as crime volume increases) and the crime market size given by the number of crime opportunities. Crimes with substantial scale economies, such as credit card fraud or vehicle theft, will be more prevalent in a few large cities, whereas other crimes with low scale economies (*i.e.*, shoplifting or minor larceny) will be prevalent in many cities of smaller size. Moreover, large cities will tend to have a wide diversification of crimes, whereas victims in small cities will tend to suffer from a smaller variety of crimes. We can already deduce that these premises will mean a distribution of urban crime that will be skewed to the right.

This connection between scale economies of urban crime and Zipf's law is based on two assumptions based on the urban scaling or urban agglomeration hypothesis (Bettencourt *et al.* 2010). The hypothesis of urban scaling predicts that certain properties of all cities change, on average, with their size in predictable scale-invariant ways (Bettencourt, Lobo, and Youn 2013). The first assumption is that crime prevalence and incidence are proportional to city size; that is, crime quantities are proportional to the number of potential victims and offenders. For instance, a city with 1 million actual crime victims might have several million more potential victims. That cannot be the case in smaller cities. Zipf's law predicts a pattern of potential victim and offender agglomerations. Second, predatory crimes, gangs, and organized crime offenders will be more prevalent in larger rather than smaller cities, as types of crimes are also a function of criminal density and proximity. With a large number of crime opportunities, the size of the crime market also increases. Likewise, the more often a person is a victim of a crime, the more likely it is that he or she will be a victim again. These particulars will tend to yield log-normal distribution of crime quantities in a ranked order system of cities.

To explain city sizes, Zipf put forward the idea that there were two main yet opposing forces in urban development that could explain the rank-size rule: the forces of diversification and unification. The first causes population dispersal and decrease, whereas the second causes population concentration and growth. The coefficient of the slope represents the ratio of the magnitude of the force of diversification to that of the force of unification; that is, the "theoretical force of diversification pulls [the slope] in the direction of a zero slope, whereas our theoretical force of unification pulls it in the direction of a true vertical, with the slope becoming infinitely great" (Zipf 1949, 131).[3] In terms of criminal victimization, the force of diversification in urban systems will increase the number of victims in cities while decreasing the number of cities, whereas the force of unification will increase the number of cities and redistribute the number of victims among them. Eaton and Eckstein (1997) identified some push and pull factors in urban economics than can help explain Zipf's law, such as distance and human capital. This is connected to the principle of least effort that is discussed in the next section. I can advance the idea that the development of the general conditions of production (*e.g.*, roads, railroads) have helped reduce transportation costs between cities and help promote production and trade.

Berry (1961) noted that underdeveloped nations were less likely to conform to Zipf's law. Systems of cities in underdeveloped nations tend to have a primate city; that is, one very large city many times larger than the next city in population size. This is not due to the degree of urbanization but to geopolitical and economic development variables. In particular, countries with a history of political colonialism or high levels of economic dependency on other countries tend to have primate cities, as is the case in Mexico. On the other hand, countries with advanced economies tend to have several large cities that are complementary rather than duplicative or parasitical (Berry 1961). Berry also noted that simple explanations cannot accurately account for differences in the types of city-size distributions. This is so, he argued, because numerous factors act randomly over a long period of time to produce a rank-size distribution in a given area (*i.e.*, imperialism, capitalism, communism, *etc.*).

It is important to note that research on the causes of Zipf's law is inconclusive. For instance, mathematical explanations cannot explain historical variations of Zipf's law when applied to aggregations of human behavior. Likewise, the mathematical explanations cannot account for the discrepancies between countries and periods of time, and cannot be accommodated to all human behaviors, as there are also exceptions to the law. There have to be other reasons that explain previous differences and deviations from Zipf's law, whether these are minor or not.

Explaining the relationship between criminal opportunity and the principle of least effort

In this section, I raise several concerns. First, there must be a sound reason why Zipf's law would be applicable to the analysis of crime distributions. Crime distributions cannot originate only and be laconically explained by the mathematical nature of power laws. There has to be an (at least one) economic variable at work behind crime quantity distributions that can help explain locational advantages for crime. Second, crime pattern theory, as a combination of opportunity theory and routine activity theory (Cohen and Felson 1979; Ratcliffe 2004), can provide theoretical and empirical support for the relationship between criminal opportunity and the principle of least effort. This relationship, at present, can only be discussed in an intuitive and hypothetical way. In the next section, I perform an empirical analysis of crime data in Mexican cities and test Zipf's law. It is not my objective in this article, however, to demonstrate this relationship, but to show the usefulness of the argument.

The relationship between crime opportunity theory and the principle of least effort seems straightforward in this respect. Criminal opportunity theory uses economic theory to predict how offenders and victims will behave, and the principle of least effort suggests that criminal activity is utilitarian and opportunistic. The relationship between the two is important for several reasons. First, according to routine activity theory, crime is contextually conditioned to three elements: a motivated offender, a suitable target, and a lack of guardianship. A crime opportunity arises every time one potential offender and one potential victim meet. An otherwise unmotivated offender is more likely to commit a crime in a place with plenty of suitable targets than one motivated offender without suitable targets. Second, from a rational perspective, offenders seek to minimize effort and maximize profits, and thus simplify their behaviors in various ways. For instance, according to the principle of least effort, the motivated offender's effort is conserved by locating a large number of suitable targets with limited or no guardianship. Third, crime pattern theory shows that criminal activity is neither spatially uniform nor random. Prior studies have explained criminal behavior based on the principle of least effort (Bernasco 2010; Vilalta 2010; Rossmo 2014). This is especially evident in the field of journey-to-crime literature (Cornish and Clarke 1987; Felson 1987; Johnson *et al.* 2007; Bernasco and Block 2009, 2011; Townsley and Sidebottom 2010; Birks, Townsley, and Stewart 2012; Baudains, Braithwaite, and Johnson 2013; Bouchard, Beauregard, and Kalacska 2013; Lammers 2014; Ackerman and Rossmo 2015; Vandeviver *et al.* 2015).

Empirical studies suggest that offenders choose crime targets close to their place of residence. Vilalta (2010), examining theft crimes in Mexico City, found that about 39 percent of convicted offenders committed their crimes in the same neighborhood where they lived, with a mean distance of 2.9 miles from the place of residence to the place of the crime and a median distance of 1 mile. Briefly put, distance to crime is a skewed distribution. As such, it can be safely assumed that criminal behavior is subject to the friction of distance (Weisburd, Morris, and Groff 2009). The principle of least effort has also been used to explain the spatial concentration of crime (Groff and Lockwood 2014; Weisburd 2015).

The relationship between crime opportunity and city size is well-established (Haynes 1973). In the United States, the dispersion of activities away from the home has had a greater impact on the crime rate in larger cities than in smaller cities (Jackson 1984). The reason is that large cities offer higher levels of anonymity and correspondingly lower levels of social cohesion and informal surveillance than smaller cities. Classic authors like Wirth (1938) and Simmel (1950) believed that social life in larger cities was fragmented and impersonal, leading to negative effects such as crime, social disorganization, and other typical urban problems such as congestion and pollution.

Cohen and Felson (1979) suggested that the convergence of factors that create opportunities for crime varies not only between neighborhoods and across time, but also among cities. Furthermore, Haynes (1973) suggested that the number and density of criminal opportunities in a city is a square function of the population size. Thus large cities have more crime victims not only because there might be more motivated offenders as a function of the city's population, but also because there are more criminal opportunities as a whole. Larger cities offer more opportunities for social interaction, which could lead to more crime opportunities (Schläpfer *et al.* 2014; Banerjee, Van Hentenryck, and Cebrian 2015). Interaction models imply that individual criminal behavior depends on individual incentives to commit a crime and on the behavior of victims. Social interactions create enough covariance across individuals to explain the high cross-city variance of crime rates (Glaeser and Sacerdote 1996).

Scale economies must also exist in the production of crime opportunities. Larger cities have a significant criminal opportunity advantage over smaller ones. Agglomeration creates a multiplier effect on the crime rate through two channels (Gaigne and Zenou 2015). First, a large city size brings greater expected monetary returns because criminals have a larger number of potential victims and a lower probability of being arrested. As Glaeser and Sacerdote (2000) argued, crime in large cities offers higher pecuniary benefits and lower probabilities of arrest due to a lower probability of recognition. Second, more people in a city increase rent and commuting costs, thereby diminishing the opportunity cost of illegal activity. This is why larger cities have higher crime prevalence, incidence, and concentration than small cities (Glaeser and Sacerdote 1999; Kahn 2010; Gaigne and Zenou 2015). This logic applies also to homicide, even though the interactions between offenders and victims happen only once (Alves *et al.* 2013). Declines in crime have been associated with city size as well (Levitt 2004).

Existing studies that have documented this correlation between city size and crime in Latin America have suggested two possible explanations for the correlation (Gaviria and Pages 2002). One explanation is that the probability of arresting a criminal is lower in larger cities, and another is that larger cities give refuge to greater proportions of crime-prone individuals. Gaviria and Pages (2002) rejected the hypothesis that larger cities have more crime because they have wealthier victims. Thus, as cities grow, and if the number of crimes and the number of police grow at the same rate, then apprehension will still be harder in larger cities because the pool of potential suspects will also be larger (Glaeser and Sacerdote 2000). In sum, in larger cities there are more crime targets, potential victims, and offenders, and the physical distance between them is less. Likewise, more commuters and longer periods of transportation increase the potential exposure to crime. In addition, larger cities might lower the costs of crime by lowering the probability of arrest and the likelihood of punishment, conditional on arrest.

Before moving on to the data and methods section of this article, it is necessary to clarify that the economic approach to crime research does not imply an economic approach to criminal policy. The economic approach to criminal policy emphasizes deterrence and incapacitation. Incapacitation might work with career criminals and deterrence might be better for amateurs or juveniles. Likewise, social crime prevention might work with both types of criminals. I find it difficult to believe, however, that crime policy should be reduced to the probability of arrest and the length of the sentence, particularly in a country like Mexico with a notable police corruption problem. I discuss this further in the discussion section and in the light of the findings of this article.

Data and methods

I test Zipf's law in thirty-two cities in Mexico using survey data from the National Survey of Victimization and Perception of Public Safety (*Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública* [ENVIPE]). This survey has been conducted by the Mexican National Institute for Statistics and Geography (INEGI) at the household level on an annual basis since 2011. It presents a representative picture of victimization at the national and state level. This survey has been conducted at the city level only twice, in 2011 and 2015. A total of sixteen cities constitute the city level sample in 2011, and thirty-two cities in 2015. A total of twelve cities were surveyed in both years (see Table 1).

The dependent variables were the number of crime victims and the number of people who reported having a fear of becoming a victim of a crime. The independent variables were the statistical rank order of the corresponding values for each data point (*i.e.*, each city). All variables were transformed for analysis using the log10 procedure. I used the ordinary least squares (OLS) method to regress the log10 transformed dependent variables on log10 scores of the statistical rank, to meet the assumptions of normality and homoscedasticity and to check for linearity of the data. In OLS regression, linearity means a constant rate of change of the dependent variable with respect to the independent variable over all values of the latter.

One might think that the previous linear correlation would provide enough evidence of Zipf's predictions. Things are not so simple, however. It might be that linearity holds but OLS results (*i.e.*, the slopes of the lines) are quantitatively different from Zipf's law. That is, the data might fit the rank size rule closely and still deviate from Zipf's equation of a descending straight line from left to right at an angle of –45°, indicating steeper or shallower slopes. It might be that the slopes differ by some factor compared to the –45° case. These deviations might reflect genuine differences between systems of cities and urbanization processes, or could be due to the lack of data of smaller units of analysis (Okuyama, Takayasu, and Takayasu 1999).

Table 1. Descriptive statistics.

Year	Variable	*n*	*M*	*SE*	Minimum	Maximum
2011	Victimization rate (%)	16	31.4%	1.3%	24.2%	47.1%
	Victims of crime	16	503,407	258,077	61,123	4,272,869
	Fear of crime rate (%)	16	80.3%	1.7%	63.3%	91.3%
	Victims of fear of crime	16	1,259,725	627,506	160,091	10,367,728
2015	Victimization rate (%)	32	31.9%	0.9%	20.7%	44.1%
	Victims of crime	32	415,792	192,878	41,518	6,279,276
	Fear of crime rate (%)	32	89.0%	0.7%	81.7%	97.8%
	Victims of fear of crime	32	1,032,031	418,063	152,201	13,626,280

For this reason, a Student's t test was used to determine whether the slopes of the regression lines (b_1) differed significantly from –1.0 (b_0).[4] If slopes are not similar (*i.e.*, if lines are not parallel), then Zipf's equation should be discarded. The absence of parallelism does not imply that the rank-size rule should be rejected, but rather that we are not in the position to corroborate Zipf's prediction that data points will invariably descend in a straight line at an angle of –45°.

Results

The unit of observation in this article is the city as defined by INEGI. In this section, I provide analysis based on the ENVIPE 2015 data set. I use these data to test the hypothesis that the distribution of crime victims and individuals afraid of crime follow Zipf's law; that is, the distribution of these two variables is characterized by a descending linear relationship of slope –1 according to their population size.

Let us begin with a descriptive look at the key variables. Although the samples are not matched, during 2011 and 2015, the victimization rate remained about the same, at 31.4 percent and 31.9 percent of the adult population. The standard errors are small. Fear of crime, on the other hand, increased substantially from around 80 percent to almost 90 percent between 2011 and 2015, suggesting that fear of crime is not only more severe than crime victimization, but it has also grown to cover almost all of the adult population in the sample. Normality tests show that the logarithmic transformations improve the data normality for both the victims of crime and those people who fear that they will be victims of crime (see Table 2). Normality tests also show that there is a very low probability that the data generating processes behind these variables are normally distributed, suggesting a concentration of crime victims and people who fear that they will be victims of crime in larger cities. Over half of all victims of crime and those with fear of becoming a victim of a crime reside in the Mexico City metropolitan area. This finding shows that Mexico is a primate system in which a few cities account for a disproportionate amount of crime victims and victims of fear of crime.

The OLS models are estimated in Table 3. Logarithmic scores not only permit valid statistical inference for skewed data, but also provide a common metric for measuring the impact of the statistical rank on both dependent variables. In OLS, the use of logarithmic scores answers if the number of victims decrease linearly to statistical rank. Let us remember that linearity means that the relationship falls on a straight line, but not necessarily in a predetermined angle of inclination. The OLS results make it clear that the linearity postulation of the rank-size rule is empirically satisfied for both the number of victims of crime and the people who fear that they will be victims of crime for both years. This implies that if the number of crime victims increases in any proportion, the number of large cities with large numbers of crime victims will decrease in a given multiple of this proportion. This finding gives some confidence to not reject the hypothesis of this article.

To get a better sense that linearity does not imply parallelism with Zipf's law, consider the scattergrams displayed in Figure 1. Coefficients suggest a superlinear (*i.e.*, more than linear) relationship between crime and statistical rank. Thus the question of whether one is able to make reasonable predictions based on Zipf's law turns to the issue of differences in the slopes of the regression lines. Aside from linearity, it is useful to know whether the regression lines follow crossing or parallel trajectories.

Table 2. Shapiro–Wilk normality tests (*z* value).

	2011	2015
Victims of crime	4.839***	6.550***
Victims of crime (log10)	2.215**	2.823***
Victims of fear of crime	4.816***	6.469***
Victims of fear of crime (log10)	2.390***	2.959***
N	16	32

*$p > 0.10$.
**$p > 0.05$.
***$p > 0.01$.

When the empirical slopes based on the data are contrasted with Zipf's theoretical slopes, the results of the tests suggest rejecting Hypothesis 2 for year 2011, because the regression lines in that year are not parallel. In fact, the slopes for the Mexican data are steeper than the theoretical slopes (see Table 4), meaning that the two lines do not have parallel trajectories as they intersect at a single point. Based on these slopes, it can be concluded that the data for 2011 show a different rate of change in angular declination from the simplest form of Zipf's law. For the 2015 data, however, agreement with Zipf's theoretical slopes is very good, particularly for those people who fear that they will be victims of crime. For 2015, the relationship between the logs of the criminological variables with statistical rank follows a straight line with a slope of near −1, corresponding to a crime victimization elasticity of statistical rank of around −1, meaning that for every decrease in statistical rank, crime victimization decreased by the same proportion.

These deviations from the simplest form of Zipf's law should not be surprising (Moreno-Sánchez, Font-Clos, and Corral 2016). Not all phenomena are statistically compatible with this form of Zipf's law. These deviations express local factors specific to individual cities (Bettencourt 2010). Likewise, it should be noted that the 2011 sample size ($n = 16$) is half that of 2015, and Zipf warned that the law only applies when there is an optimal sample size (Powers 1998).

Discussion

Because of Zipf's law, it can be stated that the number of cities with a population of crime victims greater than N is proportionate to $1/N$, meaning that (1) there is a strong log-linear relation between the number of victims in each city and its statistical rank, and (2) that a few cities tend to have the highest concentration of crime victims. This empirical regularity leads to several questions.

The first question is this: Why would crime in cities comply with Zipf's law? I believe, and will try to demonstrate with future studies, that some cities create more and better crime opportunities than

Table 3. Ordinary least squares regression results.

	2011		2015	
	Victims of crime (log10)	Victims of fear of crime (log10)	Victims of crime (log10)	Victims of fear of crime (log10)
Rank (log10)	−1.331***	−1.291***	−1.090***	−1.016***
	(0.065)	(0.080)	(0.057)	(0.054)
Intercept	6.462***	6.839***	6.485***	6.854***
	(0.058)	(0.072)	(0.067)	(0.063)
F test	415.22***	254.81***	355.87***	353.99***
R^2	0.967	0.947	0.922	0.921

Note: Standard errors are shown in parentheses.
*$p > 0.10$.
**$p > 0.05$.
***$p > 0.01$.

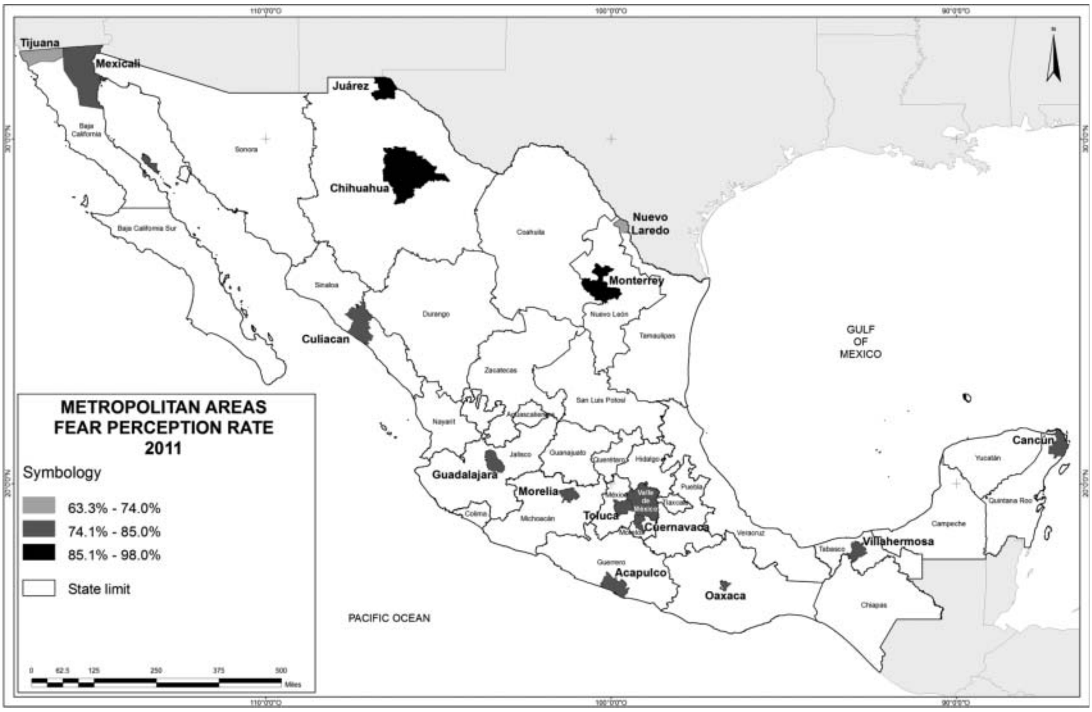

Figure 1. Map of Mexican metropolitan areas.

others, meaning that the principle of least effort also applies for the study of crime distributions in urban systems. Furthermore, I hypothesize that the general conditions of production (*i.e.*, infrastructure and equipment) also add to the explanation of crime concentrations in a small subset of cities. Simply put, the richer and the more competitive a city is, the greater the crime incidence. In this respect, one hypothesis is that crime incidence does not only correlate with the size of the city (the number of potential victims and offenders), but also with the number of crime opportunities (economic returns) proportional to size. If this is the case, crime distributions (*i.e.*, crime victimization and fear of crime) could be predictable. One practical next step would be to attempt to detect simplified yet reliable trajectories of crime victimization across cities. For instance, the use of empirically based trajectories would be relevant for the definition of priority areas for crime prevention. As a starting point, I would suggest that future political definitions of priority areas look for the places with the highest number of crime victims in general, and not only places with the higher homicide rates. Further, I suggest that the definition of priority areas should also include speed of change in prevalence rates for all

Table 4. Equality of slopes tests.

	2011		2015	
	Victims of crime	Victims of fear of crime	Victims of crime	Victims of fear of crime
b_0	−1.000	−1.000	−1.000	−1.000
Degrees angle	−45.000°	−45.000°	−45.000°	−45.000°
b_1	−1.331	−1.291	−1.090	−1.016
Degrees angle	−59.895	−58.095	−49.050	−45.720
SEb_1	−0.065	−0.080	−0.057	−0.054
t test	5.092***	3.638***	1.579	0.296

Note: Based on $n - 1$ degrees of freedom.
$^{*}p > 0.10$.
$^{**}p > 0.05$.
$^{***}p > 0.01$.

crimes. In addition, forecasting prevalence rates based on survey data would allow us to detect statistical anomalies in self-reported data from victimization surveys.

Another question is this: Why bother? Aside from ethical reasons, we should bother because crime leads to a number of social and political problems, such as public disaffection with government (Chanley 2002), community withdrawal (Hinkle 2014), greater fear of crime (Vilalta 2016), and city depopulation (Cullen and Levitt 1999). In this sense, we should bother because stakes are very high and because the empirical analysis of crime has proven to be an effective strategy to accurately quantify and prevent crime victimization for a long time now (Sherman et al. 1997; Chanley 2002; Redo 2008). Crime victims in Mexican cities today either vote with their feet, vote for anticrime policies (*i.e.*, penal populism), or choose to stay and acquire home security systems. It is in our interest to make the problem of urban crime analytically tractable and predictable, and to create reliable solutions.

Notes

1. A power law is a function with scale and base invariance, which shows a line in the log-log scale with a slope equal to a scaling exponent. Precisely, the simplest way to see a power law is to take logarithms of both sides or plot the frequency versus the statistical rank on logarithmic scales; if the data distribution follows a power law, we would see a straight line.
2. Maybe, as we will see, as the increasing returns or higher marginal profits in criminal activity might be higher in larger cities.
3. It is clear that because the cities are ranked left to right, the regression line cannot bend upward. The question, however, is why the line is not horizontal at any point, and why, in fact, the slope is near −1. The line can bend downward with any slope at any point in the statistical rank (Zipf 1949).
4. The test was the following (Paternoster *et al.* 1998): $t = \frac{b_1 - b_0}{SEb_1}$ where SEb_1 is the standard error for the coefficient of the sample data. This test provides the probability of rejecting the null hypothesis that $b_1 - b_0 = 0$.

References

Ackerman, J. M., and D. K. Rossmo. 2015. How far to travel? A multilevel analysis of the residence-to-crime distance. *Journal of Quantitative Criminology* 31 (2):237–62.

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

Banerjee, S., P. Van Hentenryck, and M. Cebrian. 2015. Competitive dynamics between criminals and law enforcement explains the super-linear scaling of crime in cities. *Palgrave Communications*, 1:15022.

Baudains, P., A. Braithwaite, and S. D. Johnson. 2013. Target choice during extreme events: A discrete spatial choice model of the 2011 London riots. *Criminology* 51 (2):251–85.

Bernasco, W. 2010. A sentimental journey to crime: Effects of residential history on crime location choice. *Criminology* 48 (2):389–416.

Bernasco, W., and R. Block. 2009. Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology* 47 (1):93–130.

———. 2011. Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *Journal of Research in Crime and Delinquency* 48 (1):33–57.

Berry, B. J. 1961. City size distributions and economic development. *Economic Development and Cultural Change*, 9 (4):573–88.

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

Bettencourt, L., J. Lobo, and H. Youn. 2013. The hypothesis of urban scaling: Formalization, implications and challenges. *arXiv Preprint arXiv* 1301:5919.

Birks, D., M. Townsley, and A. Stewart. 2012. Generative explanations of crime: Using simulation to test criminological theory. *Criminology* 50 (1):221–54.

Bouchard, M., E. Beauregard, and M. Kalacska. 2013. Journey to grow linking process to outcome in target site selection for cannabis cultivation. *Journal of Research in Crime and Delinquency* 50 (1):33–52.

Burger, M., and E. Meijers. 2012. Form follows function? Linking morphological and functional polycentricity. *Urban Studies* 49 (5):1127–49.

Chang, Y. W. 2016. Influence of the principle of least effort across disciplines. *Scientometrics* 106 (3):1117–33.

Chanley, V. A. 2002. Trust in government in the aftermath of 9/11: Determinants and consequences. *Political Psychology* 23 (3):469–83.

Chen, W. C. 1978. On Zipf's law. Doctoral dissertation, University of Michigan, Ann Arbor, MI.

Christaller, W. 1966. *Central places in southern Germany*. Englewood Cliffs, NJ: Prentice Hall.
Cohen, L. E., and M. Felson. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44 (4):588–608.
Cornish, D. B., and R. V. Clarke. 1987. Understanding crime displacement: An application of rational choice theory. *Criminology* 25 (4):933–948.
Cullen, J. B., and S. D. Levitt. 1999. Crime, urban flight, and the consequences for cities. *Review of Economics and Statistics* 81 (2):159–69.
Eaton, J., and Z. Eckstein. 1997. Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics* 27 (4):443–474.
Felson, M. 1987. Routine activities and crime prevention in the developing metropolis. *Criminology* 25 (4):911–32.
Gabaix, X. 1999. Zipf's Law and the Growth of Cities. *The American Economic Review*, 89 (2):29–132.
Gabaix, X. 1999. Zipf's law for cities: An explanation. *Quarterly Journal of Economics* 739–767.
Gaviria, A., and C. Pagés. 2002. Patterns of crime victimization in Latin American cities. *Journal of Development Economics* 67 (1):181–203.
Gaigné, C., and Y. Zenou. 2015. Agglomeration, city size and crime. *European Economic Review* 80:62–82.
Glaeser, E. L., and B. Sacerdote. 1996. Why is there more crime in cities? Working Paper No. w5430, National Bureau of Economic Research, Cambridge, MA.
———. 1999. Why is there more crime in cities? *Journal of political economy* 107 (S6):225–258.
———. 2000. The determinants of punishment: Deterrence, incapacitation and vengeance. Working Paper No. w7676, National Bureau of Economic Research, Cambridge, MA.
Groff, E. R., and B. Lockwood. 2014. Criminogenic facilities and crime across street segments in Philadelphia uncovering evidence about the spatial extent of facility influence. *Journal of Research in Crime and Delinquency* 51 (3):277–314.
Haynes, R. M. 1973. Crime rates and city size in America. *Area* 5:162–165.
Hinkle, J. C. 2014. Broken Windows Thesis. In *Encyclopedia of Criminology and Criminal Justice* 213–223. New York: Springer.
Hsu, W. T. 2008. *Central place theory and Zipf's law*. Doctoral dissertation, University of Minnesota, Minneapolis, MN.
Jackson, P. I. 1984. Opportunity and crime: A function of city size. *Sociology & Social Research* 68 (2):172–193.
Johnson, S. D., W. Bernasco, K. J. Bowers, H. Elffers, J. Ratcliffe, G. Rengert, and M. Townsley. 2007. Space–time patterns of risk: A cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology* 23 (3):201–219.
Kahn, M. E. 2010. New evidence on trends in the cost of urban agglomeration. In *Agglomeration economics*, ed. E.L. Glaeser, 339–354. Chicago: University of Chicago Press.
Lammers, M. 2014. Are arrested and non-arrested serial offenders different? A test of spatial offending patterns using DNA found at crime scenes. *Journal of Research in Crime and Delinquency* 51 (2):143–167.
Levitt, S. D. 2004. Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *The Journal of Economic Perspectives* 18 (1):163–90.
Mexico Evalua. 2015. *Prevencion del delito en Mexico: Cuales son las prioridades?* [Crime prevention in Mexico: Which are the priorities?] Mexico: Mexico Evalua.
Moreno-Sánchez, I., F. Font-Clos, and Á. Corral. 2016. Large-scale analysis of Zipf's law in English texts. *PloS ONE* 11 (1): e0147073.
Okuyama, K., M. Takayasu, and H. Takayasu. 1999. Zipf's law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications* 269 (1):125–131.
Paternoster, R., R. Brame, P. Mazerolle, and A. Piquero. 1998. Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36 (4): 859–866.
Powers, D. M. 1998. Applications and explanations of Zipf's law. In *Proceedings of the joint conferences on new methods in language processing and computational natural language learning*. In Association for Computational Linguistics, ed. D. M. Powers, 151–160. Sydney: Flinders University of South Australia.
Ratcliffe, J. H. 2004. Crime mapping and the training needs of law enforcement. *European Journal on Criminal Policy and Research* 10 (1):65–83.
Redo, S. 2008. Six United Nations guiding principles to make crime prevention work. In *International perspectives of crime prevention: Contributions from the 1st Annual International Forum*, ed. M. Coester and E. Marks. Mönchengladbach: Forum Verlag Godesburg. 5–22.
Rossmo, D. K. 2014. Geographic profiling. In *Encyclopedia of criminology and criminal justice*, ed. G. Bruinsma and D. Weisburd, 1934–1942. New York: Springer.
Rossi-Hansberg, E., and M. L. Wright. 2007. Urban structure and growth. *The Review of Economic Studies* 74 (2):597–624.
Schläpfer, M., L. M. Bettencourt, S. Grauwin, M. Raschke, R. Claxton, Z. Smoreda, G. West, and C. Ratti. 2014. The scaling of human interactions with city size. *Journal of the Royal Society Interface* 11 (98):20130789.
Sherman, L. W., D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, and S. Bushway. 1997. *Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress*. Washington, DC: U.S. Department of Justice, Office of Justice Programs.
Simmel, G. 1950. Individual and society in eighteenth and nineteenth century views of life. In *The sociology of Georg Simmel*, ed. K. H. Wolff, 58–115. Glencoe, IL: The Free Press.

Townsley, M., and A. Sidebottom. 2010. All offenders are equal, but some are more equal than others: Variation in journeys to crime between offenders. *Criminology* 48 (3):897–917.
Vandeviver, C., T. Neutens, S. Van Daele, D. Geurts, and T. Vander Beken. 2015. A discrete spatial choice model of burglary target selection at the house-level. *Applied Geography* 64:24–34.
Vilalta, C. J. 2010. Correlates of distance to crime in Mexico City. *Global Crime* 11 (3):298–313.
———. 2012. The spatial dependence of judicial data. *Applied Spatial Analysis and Policy* 5 (4):273–289.
———. 2016. Does the Mexican War on Organized crime mediate the impact of fear of crime on daily routines?". *Crime & Delinquency* 62 (11):1448–64.
Weisburd, D. 2015. The law of crime concentration and the criminology of place. *Criminology* 53 (2):133–157.
Weisburd, D., N. A. Morris, and E. R. Groff. 2009. Hot spots of juvenile crime: A longitudinal study of arrest incidents at street segments in Seattle, Washington. *Journal of Quantitative Criminology* 25 (4):443–467.
Wirth, L. 1938. Urbanism as a way of life. *American Journal of Sociology* 44 (1):1–24.
Zipf, G. K. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.