UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

HYGOR PIAGET MONTEIRO MELO

NONLINEAR SCALING IN SOCIAL PHYSICS

FORTALEZA

2017

HYGOR PIAGET MONTEIRO MELO

NONLINEAR SCALING IN SOCIAL PHYSICS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Doutor em Física. Área de Concentração: Física da Matéria Condensada.

Orientador: Prof. Dr. José Soares Andrade Junior

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M485n Melo, Hygor Piaget Monteiro.
Nonlinear scaling in social Physics / Hygor Piaget Monteiro Melo. – 2016.
66 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física , Fortaleza, 2016.
Orientação: Prof. Dr. José Soares de Andrade Júnior.

1. Sociofísica. 2. Leis de escala. 3. Eleições. 4. Alometria. I. Título.

CDD 530

HYGOR PIAGET MONTEIRO MELO

NONLINEAR SCALING IN SOCIAL PHYSICS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Doutor em Física. Área de Concentração: Física da Matéria Condensada.

Aprovada em 26/08/2016

BANCA EXAMINADORA

Prof. Dr. José Soares Andrade Junior (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. André Auto Moreira
Universidade Federal do Ceará (UFC)

Prof. Dr. Saulo Davi Soares e Reis
Universidade Federal do Ceará (UFC)

Prof. Dr. Luciano Rodrigues da Silva
Universidade Federal do Rio Grande do Norte (UFRN)

Prof. Dr. Marcelo Andrade de Filgueiras Gomes
Universidade Federal de Pernambuco (UFPE)

AGRADECIMENTOS

Meu orientador Professor Dr. José Soares de Andrade Jr., pela oportunidade, paciência e compreensão durante todo meu doutorado.

Professores Dr. André Auto Moreira e Dr. Saulo Davi Soares e Reis pela oportunidade de colaboração e ajuda indispensavéis na elaboração dessa tese.

Professores Dr. Hans Jurgen Herrmann e Dr. Hernan Makse pela oportunidade de colaboração.

Marianna Campos, por toda dedicação, companhia e paciência durante todos esses anos.

A todos os meus familiares, mas em especial aos meus pais Raimundo Rabelo Melo e Iracema Monteiro de Almeida Melo pela dedicação e apoio dados durante toda a minha vida.

A todos os meus amigos: Heitor Credidio, Diego Ximenes, Daniel Gomes, Daniel Marchesi, Davi Dantas, Diego Lucena, Diego Rabelo, Leandro Jader, Levi Leite, Rafael Alencar, Saulo Dantas, Vagner Bessa, Kauã Melo, Eduardo Araujo, Rilder Pires, César Menezes.

Todos os funcionários do departamento de Física da UFC.

As instituições UFC e ETH, assim como a FUNCAP, CAPES e CNPq pelo apoio financeiro.

RESUMO

As aplicações da mecânica estatística no estudo do comportamento humano coletivo não são uma novidade. No entanto, nas últimas décadas vimos um aumento enorme do interesse no estudo da sociedade usando a física. Nesta tese, utilizando técnicas da física, nós estudamos leis de escala não-lineares em sistemas sociais. Na primeira parte da tese realizamos a análise de dados e modelagem de eleições públicas. Mostramos que o número de votos de um candidato escala não-linearmente com o dinheiro gasto na campanha. Para nossa surpresa, a correlação revelou uma relação de escala sublinear, o que significa que o "preço"médio de um voto cresce à medida que o número de votos aumenta. Usando um modelo de campo médio descobrimos que a não-linearidade emerge da concorrência e a distribuição de votos é causalmente determinada pela distribuição do dinheiro gasto na campanha. Além disso, mostramos que o modelo é capaz de prever razoavelmente o número final de votos válidos através de um argumento heurístico simples. Por fim, apresentamos o nosso trabalho sobre alometria de indicadores sociais. Nós mostramos como homicídios, mortes em acidentes de carro e suicídios crescem com a população das cidades brasileiras. Diferentemente de homicídios (superlinear) e eventos fatais em acidentes de carro (isométrico), encontramos um comportamento sublinear entre o número de suicídios e a população de cidades, o que revela uma possível evidência de influência social na ocorrência de suicídios.

Palavras-chave: Sociofísica. Leis de escala. Eleições. Alometria.

ABSTRACT

The applications of statistical mechanics in the study of collective human behavior is not a novelty. However, in the past few decades we shaw a huge spike of interest on the study of society using physics. In this thesis we explore nonlinear scaling laws in social systems using physical techniques. First we perform data analysis and modeling applied to elections. We show that the number of votes of a candidate scales nonlinear with the money spent at the campaign. To our surprise, the correlation revealed a sublinear scaling, which means that the average "price" of one vote grows as you increase the number of votes. Using a mean-field model we find that the sublinearity emerges from the competition and the distribution of votes is causally determined by the distribution of money campaign. Moreover, we show that the model is able to reasonably predict the final number of valid votes through a simple heuristic argument. Lastly, we present our work on allometric scaling of social indicators. We show how homicides, deaths in car crashes, and suicides scales with the population of Brazilian cities. Differently from homicides (superlinear) and fatal events in car crashes (isometric), we find sublinear scaling behavior between the number of suicides and city population, which reveal a possible evidence for social influence on suicides occurrences.

Keywords: Social Physics. Scaling. Elections. Allometry. .

LIST OF TABLES

LIST OF FIGURES

5 **Analytical results of the model.** By solving our model expressed in Eq. 1, we calculate the expected number of votes of each candidate. The total number of votes divided by the number of voters n is defined as the turnout ratio T. For all 56 parliamentary elections in 2014, we compared our model estimation of the turnout ratio, T_{model}, and the data ratio T_{data}, as we see in (A). The dashed line represents what would be the perfect agreement ($T_{data} = T_{model}$) and we see that the simulations (black circles) exhibit a good agreement. We can also select the candidate with the largest number of votes $v_{\max}$ and see how our model estimate this value. In (B) we see that our model (black circles) better estimates $v_{\max}$ than the linear (red squares), which most of the time overestimates it. We show in (C) a histogram of T for the election of 2006, 2010, and 2014. We find an average turnout value of $\approx 67\%$, which is consistent with our heuristic estimation of $T = 1 - e^{-1} \approx 63\%$. We know that the exponential distribution have the property that the mean and standard deviation are equal, this property can be used in order to test if the dispersion along the mean follows an exponential distribution. In (D) we see that for state deputies of the eight largest states in 2014 election the data is in close agreement with the hypotheses of $\sigma = \langle v \rangle$. We use the exponential distribution and the expected number of votes calculate by our model to generate a random election. We show in inset, for São Paulo, that when we add the random noise to our model (red triangle) we obtain a cloud that closely resembles the actual data (black circles). p. 41

6 **Dependence with Δm.** The solution of the mean field model enables us to calculate the turnout radio T in function of the adimensional $n\Delta m/M$ parameter. In (A) we compare turnout for the approximated case where we excluded the competition between the candidates, T_{gas}, with the case with competition, T. The competition creates an exponential saturation, which increases the loss of money when candidates seek new voters. This inefficiency is maximum when $n\Delta m/M = 1.0$, we can see that by looking for $T_{gas} - T$. We can also vary Δm and see how the curve $v(m)$ change. In (B) we show that as we decrease Δm the values of $v(m)$ usually increases, as expected by the definition of Δm, however there is a point that we have a saturation process as the total number of votes starts to get close of the size of the system ($T \to 1.$). . . . . p. 44

CONTENTS

1 INTRODUCTION

Historically the origin of statistical mechanics is strongly connected with the study of statistical patterns of collective human activities. Maxwell and Boltzmann, the fathers of statistical mechanics, were greatly influenced by the emergence of new applications of probability theory on social data [1, 2]. However, it was only in the past few decades that the study of society using physics has been formalized and saw a huge spike of interest thanks to the large number of new social data produced by the Internet.

In this thesis we explore social systems using techniques commonly used in the physical sciences. The results is composed by two works [3] where we perform statistical data analysis and also a physical modeling approach. The main objective is to characterize and model the nonlinear scaling laws present at urban systems and elections.

This thesis is organized as follows: In Chapter 2, we present a short review on the introductory concepts of statistical mechanics and on its entangled history with statistics of social phenomena. As examples, we show how these concepts can be applied to two important problems: Distribution of money on a closed economical system, and the recent application of physics on the urbanization problem.

In Chapter 3, we study the scaling relation between the number of votes of a candidate and the amount of money spent at Brazilian legislative elections. Surprisingly, we find that the scaling is sublinear. This means that the price of a vote grows disproportionally with the number of votes, in such way that the richest candidates, on average, spend more for each vote than the less wealthy ones. To understand this observation we build a mean-field model that fits the relation between number of votes and money spent and allow us to explain this nonlinearity as a result of the competition among candidates.

In Chapter 4, we present our work on allometric scaling of social indicators. Allometry is a nonextensive relation between a property and the system size, often considered to follow a power-law. This nonlinear relation was first discussed by Galileo in his *Dialogues Concerning Two New Sciences*. He noted that skeletons become much more robust and

massive relative to the size of the body as the body size increases [4]. However, the most prominent allometric relation was discovered by Max Kleiber in 1947 [5], between the metabolic rate of animals and their corresponding masses. More recently those ideas were extended to urban systems by Bettencourt et al. [6]. Here we show how homicides, deaths in car crashes, and suicides scales with the population of Brazilian cities. Our findings support the hypothesis that the number of suicides may be influenced by the non-trivial social substrate of cities.

Finally, in Chapter 5, we present our conclusions and explore the perspectives for future works.

2 BASIC CONCEPTS OF STATISTICAL MECHANICS

Dealing with uncertainty is an unavoidable consequence of human nature. This happens because science has the objective to explain, describe and predict natural phenomena and this always occurs under the state of reasoning with conditions of incomplete information. This lack of information happens in many levels and for many different reasons. Even if you take a simple length measurement with a rule, we know that different results will emerge in each try. This universal lack of information is represented in scientific theories by randomness, and this universal aspect of nature is what gives the actual importance of statistics and probability theory.

Statistics originated in 17th century, motivated by the need to make sense of the social numbers collected by the increasingly bureaucratic state machine, such as the rates of death, birth, and marriage. The term statistics was introduced in the 18th century to denote these studies dealing with civil "states" [1, 2].

Inspired by the great success of Newton's mechanics and astronomy, some scientists and philosophers start to seek immutable "natural" laws that governed human society. The French astronomer Pierre-Simon Laplace showed that the variations in male and female births and other social statistics could be described by a universal law, which we know today as Gaussian or normal distribution. This law was first proposed to describe the probabilities of coin tossing, which led Laplace to conclude that the birth of a male or female is a result of a random process, not a God's desire to provide spouses for all. The success of Laplace's studies on social data with the astounding ubiquity of the Gaussian curve of errors led the astronomer Adolphe Quetelet to write that

> "whatever concerns the human species, considered *en masse*, belongs to the domain of physical facts; the greater the number of individuals, the more the individual will is submerged beneath the series of general facts which depend on the general causes according to which society exists and is conserved [7]."

Statistical mechanics first appeared to deal with a specific type of uncertainty generated by insufficient computational power. Before the discovery of quantum mechanics, physics was a deterministic science. However when you take a system with an enormous number of elements, there is a fundamental impossibility to solve all equations of motion. Even if it was possible, the amount of data generated would be intractable. Remarkably, this lack of information can be used as an advantage, the application of statistics on physical system with huge number of elements shows that we can ignore most of the microscopic rules of interaction and, from an apparent microscopic randomness, emerges a homogeneous system governed by simple mathematical laws. This new physical methodology is explained by James Clerk Maxwell when he came to study the problem of gases:

> "...the smallest portion of matter which we can subject to experiment consists of millions of molecules, not one of which ever becomes individually sensible to us. We cannot, therefore, ascertain the actual motions of any one of these molecules; so that we are obliged to abandon the strict historical method, and to adopt the statistical method of dealing with large groups of molecules... In studying the relations between quantities of this kind, we meet with a new kind of regularity, the regularity of averages, which we can depend upon quite sufficiently for all practical purposes [8]."

The first contributions on the formal development of statistical mechanics were made by Daniel Bernoulli, James Clerk Maxwell, Ludwig Boltzmann, and Josiah Willard Gibbs in the second half of the nineteenth century. They were faced with the problem of explaining the empirical laws of thermodynamics by using the principles of Newtonian physics. To do so, they had to assume the existence of atoms and make extensive use of probability theory, a mathematical branch that can be traced back to mathematicians interested in maximizing their profit in games of change. These two ingredients produced a deep impact in science, initially with the creation of statistical physics and later with the application of this conceptual framework to interdisciplinary fields like biology, sociology, economy, etc [9, 10, 11, 12].

Even though one might think that Social Physics is a new endeavor, the two areas have intermingled and influenced each other since the times of Maxwell, as put by the science journalist Philip Ball [1]:

> "Contemporary efforts to apply the concepts and methods of statistical physics to social phenomena ranging from economics to traffic flow, pedestrian mo-

> tion, decision making, voting and contact networks are therefore essentially completing a circle whose trajectory commenced centuries previously. Work on social statistics in the 19th century had a direct influence on the founders of statistical physics, who found within it the confidence to abandon a strict Newtonian determinism and instead to trust to a 'law of large numbers' in dealing with innumerable particles whose individual behaviors were wholly inscrutable."

2.1 Statistical Physics

We mentioned that the creation of Statistical Mechanics was an effort to explain from first principles the origin of the experimental results of the Thermodynamics. This achievement came through by using two ingredients: the atomic hypotheses and the use of probability theory. Indeed, theses two ideas are close connected, because the macroscopic properties of the system can be computed by assigning a probability for each of the possible states accessible to the very large number of particles that compose it, without the need to solve their equations of motion.

A typical thermodynamical system is a gas. For instance, one liter of oxygen at standard temperature and pressure ($T = 273,15K$ and $p = 1.0Atm$) consists of about 3×10^{22} oxygen molecules, all interacting with each other and colliding with the walls of the recipient [13, 14]. Although the exact state of each atom is chaotic and unpredictable, the thermodynamic quantities, like volume or pressure, follow simple relations like Boyle's law:

$$PV = constant. \tag{2.1}$$

This simplification happens because quantities are actually averages of microscopic states, where fluctuations tend to cancel out each other. Statistical Mechanics identify this average nature of the macroscopic world without trying to solve exactly the equations of motion. Instead, it proposes a method to determine the probabilities of finding the system in each microscopic state and, by evaluating the moments of this probability distribution, we can find the thermodynamic macroscopic variables. As an example, the Temperature is proportional to the first moment (average) of kinetic energy distribution. Therefore, Statistical Mechanics change the problem from solving many equations of motion to the problem of finding a suitable distribution of probability over possible states. There are

some strategies to determine such distributions over states, as we are going to describe next.

2.1.1 Master Equation

Suppose that we have a set of different systems, each one of many different random states. If we observe how they jump from one state to another, after some time we may see that the majority of systems would have converged to the same, most likely, state. This description is formalized by the method of solving a *master equation*.

Assume that our system is in a state i. The probability to find our system in a state j after a time interval of dt is defined as $R_{ij}dt$, where R_{ij} is know as the transition rate from i to j. Since we are interested in calculating which state is more likely, we have to define the probability of finding the system at state i at time t as $p_i(t)$. Therefore, we can use the transition rates to write down a master equation for the temporal evolution of $p_i(t)$:

$$\frac{dp_i}{dt} = \sum_j [p_j(t)R_{ji} - p_i(t)R_{ij}]. \tag{2.2}$$

The positive term represents the influx, the rate at which the systems are reaching state i from every other state j, and the negative term is the outflux, the rate that the dynamics remove systems from state i. The set of probabilities also have to obey a normalization condition:

$$\sum_i p_i(t) = 1. \tag{2.3}$$

The solution of Equation 2.2 with the normalization constraint gives us a method to find the distribution p_i over time. The connection with the macroscopic reality is done by taking averages with p_i. Suppose that you are interested in a macroscopic property M, if this property for each state i assumes the value of M_i, then the expected value of M at time t is

$$\langle M \rangle = \sum_i M_i p_i(t). \tag{2.4}$$

2.1.2 Equilibium

We can now ask ourselves what happens to the system when we wait enough time. If we look Equation 2.2 we can see that, if for every j the two terms on the right-hand side cancel one another, the probability distribution p_i stops to evolve ($dp_i/dt = 0$) and the system reaches a state called equilibrium state. Since equilibrium statistical mechanics is concerned only with the equilibrium state of systems, from now on we will write p_i as a shorthand for $\lim_{t\to\infty} p_i(t)$.

For a system in thermal equilibrium with a reservoir at temperature T, Gibbs showed that the equilibrium distribution probability is

$$p_i = \frac{e^{-\epsilon_i/kT}}{Z} = \frac{e^{-\beta\epsilon_i}}{Z}. \tag{2.5}$$

This is know as the Boltzmann-Gibbs distribution [14], where ϵ_i is the energy of state i and $k = 1.38 \times 10^{-23} J/K$ is the Boltzmann's constant. Using the convention $\beta = 1/kT$ we can write the normalization factor Z as

$$Z = \sum_i e^{-\beta\epsilon_i}. \tag{2.6}$$

The normalization factor Z is know as the *partition function*, and it has great significance on the mathematical treatment of statistical mechanics. The study of the variation of Z with parameters of the system, like temperature, gives us informations about the macroscopic properties.

From the Boltzmann-Gibbs distribution, Equation 2.5 together with Equation 2.4, we can calculate any macroscopic variable, for example, the expectation value of the energy $\langle E \rangle$, which in thermodynamics is defined as the internal energy U

$$U = \langle E \rangle = \frac{1}{Z} \sum_i \epsilon_i e^{-\beta\epsilon_i}, \tag{2.7}$$

and can also be expressed in function of the derivative of partition function with respect of β

$$U = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \tag{2.8}$$

From thermodynamics we know that the specific heat is given by the derivative of the

internal energy, so we have

$$C = \frac{\partial U}{\partial T} = -k\beta^2 \frac{\partial U}{\partial \beta} = k\beta^2 \frac{\partial^2 \ln Z}{\partial \beta^2}. \tag{2.9}$$

Basically the same procedure can be done to find any other thermodynamic parameter as a function of Z.

As we mentioned before, there are multiplies ways to derive the Boltzmann-Gibbs distribution. One of the most simple demonstrations is by dividing a system into two parts. Since the energy is additive, the total energy is the sum of the parts: $\epsilon = \epsilon_1 + \epsilon_2$, whereas the probability is given by the product rule: $p(\epsilon) = p(\epsilon_1)p(\epsilon_2)$. The solution for these equations is an exponential distribution, like in Equation 2.5.

Another derivation, proposed by Boltzmann, is a bit more complicated, but introduces new ideas that have a profound importance on the derivation of others equilibrium distributions. Take a system composed of N particles, where each particle has an energy ϵ_i. We call n_i the number of particles in a single state ϵ_i, such that $N = \sum_i n_i$ and the total energy $E = \sum_i n_i \epsilon_i$. Consider that the system has a finite number of energy states with $i \in [0, m]$. The number of possible configurations Ω available for a fixed E that follows these two constraints is given by the combinatorial formula

$$\Omega(E) = \frac{N!}{n_1! n_2! n_3! ... n_m!}. \tag{2.10}$$

Assuming that every configuration $\{n_i\}$ has the same probability, to find the equilibrium configuration we just need to find which E maximizes Ω, which should be the most likely state. Boltzmann noted that instead of maximizing Ω, it would be much easier to maximize $\ln(\Omega)$, because in the limit of large N we can use the Stirling approximation, $n_i! \approx e^{-n_i} n_i^{n_i}$, so

$$\frac{\ln \Omega}{N} = -\sum_i \frac{n_i}{N} \ln\left(\frac{n_i}{N}\right). \tag{2.11}$$

The problem of maximization with constraints can be easily solved by using the method of Lagrange multipliers [14]. Looking at Equation 2.11 Boltzmann noted that $\ln(\Omega)$ have an important role on statistical mechanics, in fact he defined $S = k \ln(\Omega)$ as the microscopic interpretation of Clausius' entropy [14]. Therefore, to find the equilibrium distribution, we have to maximize the Entropy, $S = k \ln(\Omega)$. This result elegantly connects

statistical physics and the second law of thermodynamics.

Equation 2.11 also give us another definition for Entropy, that was first discovered by Shannon when working with information theory [15]. Since the probability is defined as $p_i = n_i/N$, we see that another equivalent definition for Entropy is $S = -\sum_i p_i \log(p_i)$. The Shannon information entropy connected the information theory with statistical physics. This remarkable theory was responsible for the creation of a number of new interdisciplinary applications for statistical physics [10, 11, 16].

2.2 Social physics

The French political philosopher Auguste Comte in the 19th century first coined the term *Social Physics* [17, 1]. He believe that this new discipline should be the natural evolution in a search for a complete description of the world,

> "Now that the human mind has grasped celestial and terrestrial physics, mechanical and chemical, organic physics, both vegetable and animal, there remains one science, to fill up the series of sciences of observation – social physics. This is what men have now most need of; and this is the principal aim of the present work establish."

As we mentioned before, the emergence of a Social Physics was a natural consequence of the creation of Statistical Physics. Unfortunately, this historical connection was lost for the most part of the 20th century. However, in the 90s, as the Statistical Physics was already well tested and matured, there was finally a movement towards applying its concepts to sociology and economy, mostly driven by the rapid evolution of computers and the huge amount of new social data available.

Applying statistical physics to model social dynamics involves two challenges. The first is finding a simple yet comprehensive microscopic model, and the second is finding how to extract the macroscopic behavior out of it. Solving the second problem is where lies the essence of what statistical physics has to offer to social science. A common mistake is to assume that the macroscopic behavior is a straightforward extrapolation of the individuals. The political scientist Michael Lind explains this problem [18]

> "A friend of mine who raises dogs tells me that you cannot understand them unless you have half a dozen or more. The behavior of dogs, when assembled

> in sufficient numbers, undergoes an astonishing change. They instinctively form a disciplined pack. Traditional political philosophers have been in the position of students of canine behavior who have observed only individual pet dogs."

Despite of all complexity of the human interactions, the large-scale behavior seems to be independent of individual social characteristics, allowing us to use the same techniques to study of systems composed by a huge number of atoms or molecules [10]. In the following sections, we shall explore some examples of social phenomena that have been studied using the techniques of Statistical Physics.

2.2.1 Statistical Mechanics of Money Distribution

We saw that the fundamental law of equilibrium statistical mechanics is the Boltzmann-Gibbs law for the probability distribution of energy ϵ

$$P(\epsilon) = \frac{e^{-\epsilon/kT}}{Z}. \tag{2.12}$$

The analytical derivation of an equilibrium distribution needs the imposition of constraints, which usually is a conserved physical quantity, for instance, for the case of a Boltzmann-Gibbs distribution, we have to impose energy conservation.

Dragulescu and Yakovenko [19] took this classical mechanical assumption and applied it for a system of many economical agents that interact dynamically exchanging money, but the total amount of money is conserved. For simplicity, they assumed a closed system where the total amount of money is conserved, therefore, the equilibrium probability distribution of money $P(m)$ is a Boltzmann-Gibbs distribution.

Let us suppose that we have a system with a constant number of agents and each agent i has some money m_i from which a fraction can be exchange with other agents. First assume the simple case for the interaction of i with j, where the first pays an amount of money Δm for some goods or services from j. The money balance should be written as

$$\begin{aligned} m_i &\to m_i' = m_i - \Delta m, \\ m_j &\to m_j' = m_j + \Delta m. \end{aligned} \tag{2.13}$$

Naturally, this interaction rule conserves the total amount of money

$$m_i + m_j = m'_i + m'_j. \tag{2.14}$$

This local conservation is analogous to the transfer of energy from one molecule to another in molecular collisions of a gas.

Assuming money conservation and following the Boltzmann formalism, the distribution of money m is given by the exponential [19]

$$P(m) = \frac{e^{-m/T}}{Z}, \tag{2.15}$$

where the partition function $Z = \int_0^\infty e^{-m/T} dm = T$, and T is an effective temperature, or *money temperature*, equal to the average amount of money per economic agent,

$$T = \langle m \rangle = M/N. \tag{2.16}$$

Here, M is the total money, and N is the number of agents.

It is possible to perform a simulation in order to verify if Equation 2.15 holds. If you give an initial amount of \$1000 for all agents and at each time step you choose randomly two agents to perform an exchange of money Δm, as shown in Equation 2.13, after a transient period, the money distribution should converge to the Boltzmann-Gibbs distribution. The simulated money distribution and theoretical curve are plotted in Figure 1. There is experimental evidence that the wealth and income distribution for the majority of the population indeed show this exponential distribution but followed by a power-law decay [11].

This simple example shows how to apply the concepts of statistical physics for a social system where the agents are exchanging not energy, but money. The application of physical concepts and techniques to Economics represents a vast and promising field, ranging from studying correlations in stock market series [20, 21], wealth distribution [22, 11], to the analysis of income distribution and inequality [23].

2.2.2 Nonlinear Scaling

By now, we should have made clear that statistical mechanics has the objective to infer the macroscopic properties of a system from its microscopic dynamical rules. An important property to note is that, when looking at a finite system, we automatically associate to it a system size. The size may refer to the volume of the system V, or the

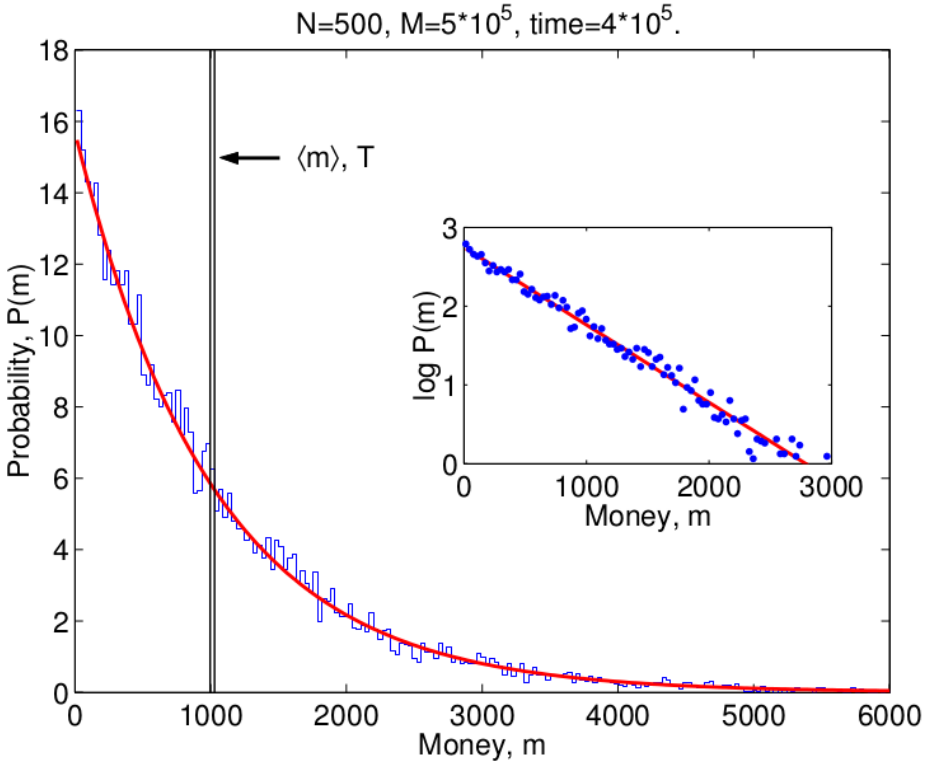

Figura 1: **Histogram of money distribution.** The blue line and points shows that the Boltzmann-Gibbs distribution emerges from a simulation of a economic model with money conservation. The vertical line shows the initial delta distribution of money. We see in the solid red line that an exponential function $exp(-m/T)$ fits the data, where the money "temperature" is $T = M/N$, with the total money $M = 5\ .\ 10^5$ and the number of agents $N = 500$. This figure was taken from Ref. [19].

number of agents N, usually persons for the case of social physics applications. Correctly assigning a size for a system is essential because comparing macroscopic variables of systems with different sizes can gives us important informations about how the agents interact.

In statistical physics we see that some measurable properties typically scale in proportion to the system size. An example is the energy, if we double the number of particles, the energy doubles as well. We call such physical properties as extensive. This means that if we have two systems A and B, when we compose those two systems any extensive property should mathematically behave as $f(A+B) = f(A)+f(B)$. Therefore, for those physical properties the whole is the sum of the parts.

Interestingly, proportional scaling is not the only way that a variable can scale with the system size. Indeed, it is not uncommon in nature to observe properties that present non-trivial forms of scale dependence. In statistical physics, for instance, when a system

is close to a phase transition, properties below the critical point are extensive, scale nonlinearly with the system size, more precisely following a power law. An example is the scaling between

$$c \propto L^{\beta}, \tag{2.17}$$

where c is the specific heat, L is the system size, and β is called a critical exponent.

In biology, the nonlinear scaling is called allometry. Allometry often implies the use of power-laws to describe the dependence of a wide range of anatomical, physiological and behavioral properties. Precisely, if we denoted Y a biological property and M the body mass of different animal species, one may find a relation in the form of $Y \propto M^{\beta}$. One of the most prominent allometric relations found in natural sciences is the so-called "three-quarters law" or Kleiber's law [5], where the metabolic rate of animals should scale with exponent $\beta = 3/4$ of their masses.

Inspired by biological laws, a group of physicists [6] showed that the cities belonging to the same urban system exhibit nonextensive rates of innovation, wealth creation, patterns of consumption, human social behavior, and several other properties related to the urban infrastructure. However, unlike in biology, they found that the scaling can be superlinear ($\beta > 1$), linear ($\beta = 1$), and sublinear ($\beta < 1$). Bettencourt and collaborators proposed a classification for the social variable in terms of the exponent [6, 24, 25]:

1. ($\beta < 1$) Sublinear scaling implies an economy of scale, because its per capita measurement decreases with population size. In the case of cities, for example, we see this type of scaling in the number of gasoline stations, the total length of electrical cables, and the road surface (material and infrastructure).

2. ($\beta = 1$) The linear (isometric) case typically reflects the scaling of individual human needs, like the number of jobs, houses, and water consumption.

3. ($\beta > 1$) Superlinear scaling in urban indicators emerges whenever the complex patterns of social activity have significant influence. Wages, income, disease, growth domestic product, bank deposits, as well as rates of invention, measured by new patents.

An explanation for the presence of allometry has been the source of long discussion [24]. The most accepted theory proposes that the 3/4 exponent in the metabolic scaling of animals comes from a transport theory applied in a fractal geometry in order to simulate the circulatory systems of animals. In the context of social physics, equivalent theories have been proposed to explain the allometric scaling in cities. Bettencourt has proposed a mean field model that uses the fractal nature of cities [25]. However, there is

still no consensus on the origin of nonlinear scaling in biology and sociology [26, 27].

3 THE ECONOMY OF ELECTIONS

3.1 Introduction

Free and fair elections play a central role in democracy [28], exhibiting a complex process of negotiations between politicians and voters, with the past decades bearing witness to a steep increase in the expenditure of political campaigns. Take the example of the presidential elections in the US. The 1996 campaigns cost contestants approximately $123 million (corrected for inflation) all together, an amount that escalated to nearly $2 billion in 2012 [29]. Although campaign investments have grown, the impact of money into the electoral outcome remains not fully understood [30], and conclusions about the theme are quite contradictory. Some studies argue that incumbent spending is ineffective, and the challenger spending, on the other hand, produces large gains [31]. Others claim that neither incumbent nor challenger spending makes any appreciable difference [32], a theory that dates back to the 1940's [33]. Yet another group argues that both challenger and incumbent spending are effective [34].

Despite the questioning about the effectiveness of political campaigns as a whole, the election campaign of President Barack Obama in 2012 spent more than 65% of its money on media, including TV and radio air time, digital and printing advertising, and others [35]. Therefore, the direct contact with voters not only figures as a major factor in campaign planning, but it is clearly believed to have relevant impact in succeeding to persuade undecided voters [36].

We try to address the problem of how campaign expenditure influences election outcome by developing a mean field model for the negotiations between candidates and voters. This model has two simple ingredients: the number of votes candidate i receives is limited by his/her money m_i, and candidates can openly compete for voters. By taking into account the competition between candidates when convincing undecided voters, the theory

accurately reproduces the vote distribution $P(v)$ when provided their money campaign distribution $P(m)$. The model shows that the top vote-receivers are those who spend more in political campaign, but with a highly counterintuitive result: the more candidate spends, higher the price of a vote.

In Economics, a similar effect in which larger companies tend to produce goods at increased per-unit costs is known as diseconomy of scale, and it may have diverse causes like communication costs [37], and the Ringelmann effect, the psychological effect when individuals get less efficient when working in larger groups [38].

This diseconomy of scale is present on real data acquired from recent proportional elections in Brazil, openly available in [39]. As far as we know, this is the first work to report such phenomena in elections. The data is related to the elections for the national lower house and state congress in 2014. Brazilian representative elections are an excellent experiment due to a number of special factors. First, Brazil is a large country, both in population and land area. It has the fifth population of the world spread across roughly 8.5 million km^2 (over 3 million mi^2). Second, in contrast with executive elections, representative elections in Brazil have a large number of candidates. Additionally, it is compulsory to vote in Brazil. Altogether, these factors leads to a huge data set from a quite diverse electorate.

We argue that competition between candidates is the root cause of this verified diseconomy of scale in Brazilian elections, showing that a scenario without competition always overestimate the number of votes of top campaign spenders. Moreover, we show that the correct estimation of the approach with competition is able to reasonably predict the turnout rate of elections through a simple heuristic argument. Finally, we test our theory by using the principle of maximum entropy to show that the statistical dispersion of the model agrees with the one found in our data set.

3.2 Materials and Methods

Here, we are able to investigate the effect of the investment of candidates on campaign thanks to the recent publication of data containing the total donations and expenses for each candidate. We analyze Brazilian elections for two different kinds legislators, federal and state deputies. The data are available at the website of the Brazilian Federal Electoral Court [39]. By force of law each candidate must provide a detailed description of each personal spending, with many specific informations like the value, date and type of

expense. All these informations can be accessed by the public, however in order to know the total cost of the campaign and the number of votes of each candidate is necessary to process the database computationally.

State deputies are local representatives elected for a term of four years by a proportional system. His function is to legislate in the unicameral system of each state. The federal deputies are representatives in the chamber of deputies of the national Congress, one of the two houses of the legislative power. The number of elected federal deputies is proportional to the population of the state, from a total of 26 states and a federal district.

3.3 Empirical findings.

We start our analysis by assembling the data sets on the entire electoral outcome and campaign expenditure of candidates from all 26 Brazilian states. Figure 2 displays the number of votes v versus the declared campaign expenditure m of each candidate in the top 5 Brazilian states by population: São Paulo (Figs. 2A and 2B), Rio de Janeiro (Figs. 2C and 2D), Minas Gerais (Figs. 2E and 2F), Bahia (Figs. 2G and 2H), and Rio Grande do Sul (Figs. 2I and 2J). Although spread, the clouds of points are neatly correlated and follow a clear trend. Remarkably, this trend is observed in all representative elections for all Brazilian states.

To extract the main relationship between v and m, we average the number of votes in log-spaced bins along m, which provides an estimation for the empirical relation of v as a function of m. In order to plot in the same figure these results for different states, we perform a scale transformation on v by supposing simple linear relation $v = c \times m$, where c is a characteristic constant of a given election. If we define the average price of a vote as $\Delta m = \sum_i m_i / \sum_i v_i$ and suppose that it is roughly uniform across candidates, it is easy to see that $c = 1/\Delta m$. If the relation between votes and money is linear then the plot of $v \times (\Delta m/m)$ should be a constant function of m with value close to 1.0.

In Figure 2, we plot $v\Delta m/m$ as a function of m for the (K) state legislative assembly and (L) federal congress elections in the year of 2014 for the eight most populated states in Brazil. The result shows a consistent nontrivial dependence of votes on money. For small values of m we see a rapidly decreasing of $v\Delta m/m$. When we look at this region in Fig. 2 we see no clear correlation between v and m, meaning that the money is so little that the votes are nearly random. Looking at the region above ten thousand Brazilian reais (R$10,000.00), we see a region with an apparent linear dependence of v with respect of

m. However, when as m increases, a noticeable departure from linearity is observed. The wealthier candidates display a lower fraction of votes per money. This sublinear scaling means economically that the candidates with more votes spend more money per vote if compared with less successful candidates and is our first main result.

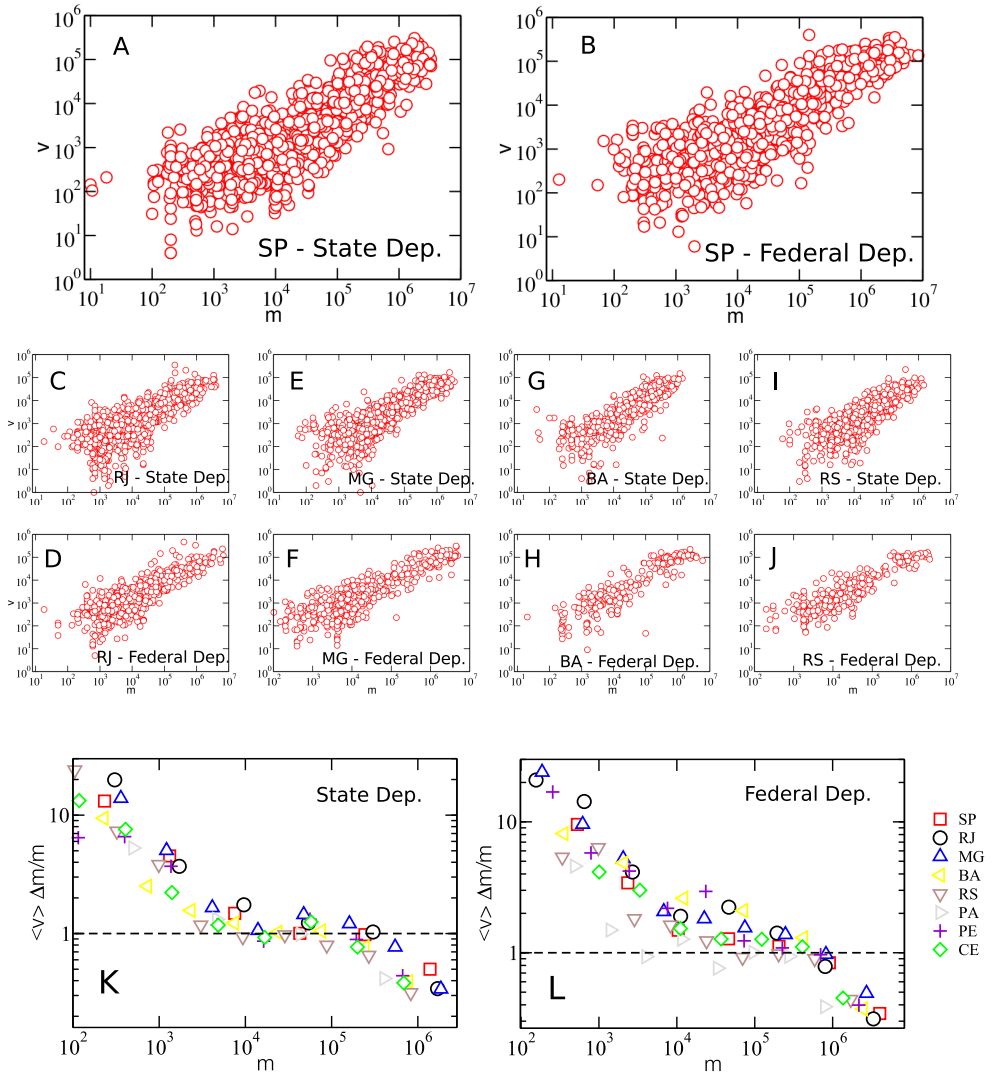

Figura 2: **Scaling Relation between number of votes and money spent.** The red circles shows the relation between the number of votes and the declared campaign expenditure of each candidate in the state and federal deputies elections in 2014 for the five largest states in Brazil: São Paulo (A,B), Rio de Janeiro (C,D), Minas Gerais (E,F), Bahia (G,H), and Rio Grande do Sul (I,J). Despite the large fluctuations, there is an unambiguous correlation between votes and money. In order to see the nuances of the correlation we plotted in (K) and (L) a normalized relation for state and federal deputies for the eight largest states in Brazil. The symbols represents the normalized ratio $\langle v \rangle \Delta m/m$ where we first calculate the average number of votes in log-spaced bins along m. If we assume a linear correlation, the multiplicative constant is $\Delta m = M/n$. The normalization provides us a direct observation of the nonlinearity in the dependence of votes on money. We see a global sublinear behavior, where the wealthier candidates display a lower fraction of votes per money.

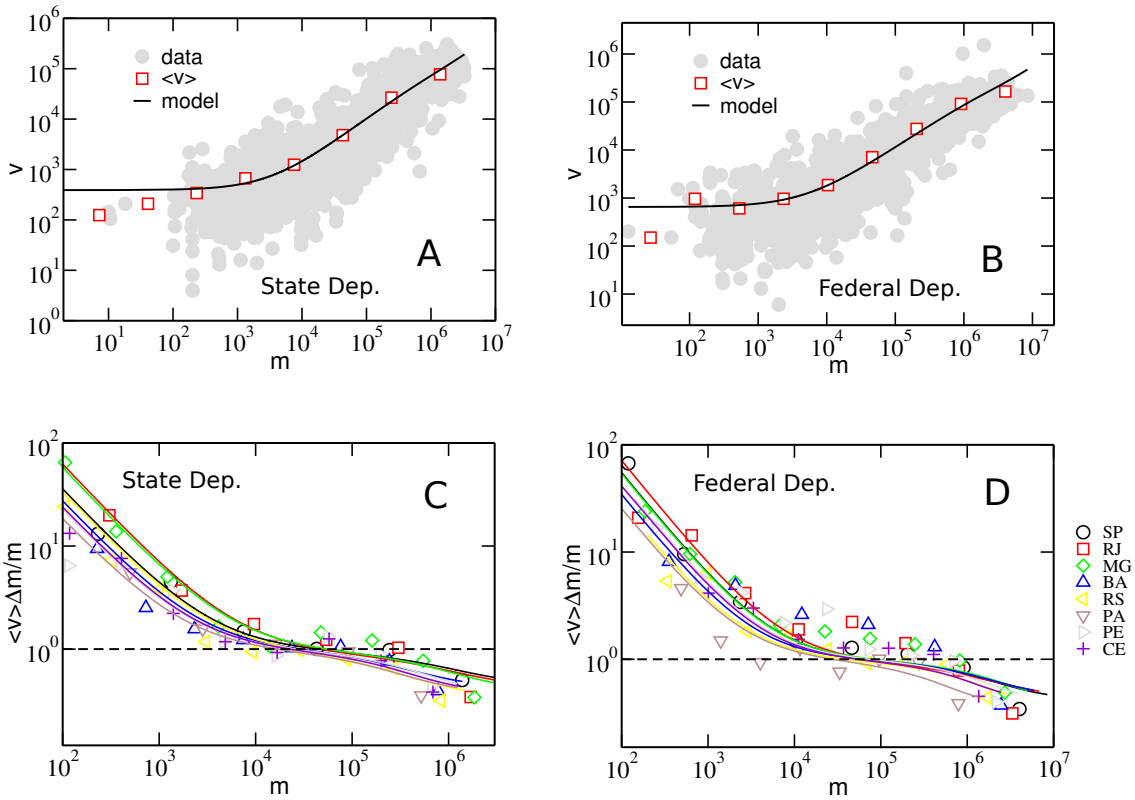

Figura 3: **Modeling the nonlinear scaling.** In order to verify if our model correctly fits the data, we show in (A) and (B) the São Paulo election for state and federal deputies in 2014, respectively. Each small circle is one candidate and the red squares are the average number of votes in log-spaced bins along m. We see that our model shows a good agreement with the average behavior for all the money spectrum. In (C) and (D) we perform the same normalization process as in Fig.1, but now with Δm estimated using Eq. 3.3. Each solid line shows the solution of our model, and the color indicates the state. Despite its simplicity, our model features all nonlinear regimes seen in the data, which corroborates our theory that the inefficiency of wealthier candidates are due mainly to competition.

3.4 A mean field approach for the price of a vote

We consider an electoral process composed of two separate groups of individuals, candidates and voters. All s candidates can compete for the vote of all n voters, and each candidate i has a limited money m_i to spend on their campaign. Thus, if at a given time m_i becomes null, the candidate will not be able to compete for voters anymore, so their number of votes v_i stops varying.

Limiting the candidates to conquer a single vote at a time and assuming that voters

already convinced cannot change their minds, we can write the time evolution of the number of votes of a given candidate i as

$$\frac{dv_i}{dt} = \left(1 - \frac{S(t)}{n}\right)[m_i(t) > 0]. \tag{3.1}$$

Where $S(t) = \sum_{i=0}^{s} v_i$ is the total number of decided voters at time t, and $[m_i(t) > 0]$ is the Iverson bracket, which is 1 if the condition inside the brackets is satisfied, and 0 otherwise. The right-hand side of the Eq. 3.1 is the probability of candidate i to choose an undecided voter at time t. Eq. 3.1 explicitly requires a definition for the rate of money expenditure dm_i/dt, in order to determine the end of the monetary resources of a candidate. Due to the nature of the problem, it is natural to expect that $dm_i/dt < 0$. We choose the simplest case where the money decreases linearly at a constant ratio, in other words $dm_i/dt = -\Delta m$. Additionally, we consider this ratio as uniform for all candidates.

The probabilistic feature of Eq. 3.1 is central to confirm our hypothesis that electoral outcome is an output of campaign expenditure due to a competition process. To show that, we first consider the case without competition, where $s \ll n$. Also, we assume here that $n\Delta m \gg m_i$ for all i, so that the candidate with the largest money do not have enough money to reach out the whole network. By doing so, it is unlikely to a candidate to waste their campaign money on voter who's already decided. Thus, the decided voters of other candidates do not affect the probability of candidate i to conquer an undecided voter, so v_i replaces $S(t)$ in Eq. 3.1, and the system of differential equations become uncoupled. Under this scenario, these equations are easily integrated, and the number of votes becomes $v_i = n - (n - v_{0,i})e^{-m_i/n\Delta m}$. Here, $v_{0,i}$ is the initial number of votes of candidate i. Since $n\Delta m \gg m_i$, and assuming $(n - v_{0,i}) \approx n$, we find the linear form of $v_i \approx v_{0,i} + m_i/\Delta m$ by expanding the exponential function and taking its first order approximation. As we will discuss next, this simple model do not suffice to explain the whole complexity of the relation between v and m.

The first two regimes presented in Fig. 2 are understood by the analytical approach form above. For the regime of low m, where the experimental data do not exhibit a clear correlation, the candidates start the race with $v_{0,i}$ votes. Since they cannot afford a long run and/or a large expenditure, their final performance fluctuates around the initial value $v_{0,i}$, which depends on different factors, such as free volunteer engagement. As campaign money increases, the linear part overcomes the initial $v_{0,i}$, and a linear regime emerges. However, the scenario without competition the linear behavior remains at large m. We shall consider the competition between candidates as a possible cause for the transition

from the linear to the sublinear regime. This is done by solving Eq. 3.1 in its entirety. By integrating Eq. 3.1, we find

$$v_i = v_{0,i} + \frac{m_i}{\Delta m} - \frac{1}{n\Delta m}\int_0^{m_i} S(m')dm', \tag{3.2}$$

where we identify that integrating the Iverson bracket over time we get the time length candidate i has to campaign, which is given by $m_i/\Delta m$. Additionally, we changed the variable of integration on the last term by remembering that $dm'/dt' = -\Delta m$.

It is possible to find a differential equation for $S(m')$ by taking Eq. 3.1 and summing over i. After solving it for $S(m')$ and integrating the last term of Eqs. 3.2 (detailed solution can be seen at section 3.7), we find a set of nonlinear coupled equations that must be solved in a specific order, candidate by candidate, following an increasing order of m_i. As a consequence, the number of votes of candidate i depends neither solely on m_i nor on t, but on the whole distribution $P(m)$ through the integral at the right-hand of Eqs. 3.2.

Equation 3.2 has a simple interpretation. As in the scenario without competition all candidates begin their run with an initial number of votes, and those with sufficient money to keep running enter in a linear regime controlled by the second term of Eqs 3.2. Nonetheless, as we will see next, candidates with large enough money may start to waste their money on decided voters, fact expressed by the presence of $S(m')$ on the last term of Eqs. 3.2, which encloses the competition dynamics. This collective influence of the campaign money from all candidates, Eqs. 3.2, is our second and most important result. It provides a bridge between campaign expenditure and electoral outcome and is the basis of the remaining results that follows.

In order to solve the model we have to use as inputs the money m of each candidate, an initial number of votes v_0, the total number of voters n and Δm. To define v_0 we use the average of votes for candidates with less then $R\$1,000$. For our model Δm behaves as a free parameter that can be calculated in function of the turnout rate $T = S_f/n$, or final fraction of votes. This is possible because the total number of votes $S(t)$ have a simple solution (see section 3.7) for the limit $t \to \infty$, and we can write the final fraction of votes as

$$T = 1 - e^{-M/(n\Delta m)}. \tag{3.3}$$

Where $M = \sum_i m_i$ is the total money. Therefore, we estimated Δm using Eq. 3.3 such that the total number of votes fits the turnout election data.

The São Paulo election for state and federal deputies in 2014 are shown in Figure 3(A) and (B), respectively. Each small circle represents one candidate and the red squares are the average of votes in log-spaced bins along m. We see that our model, the black line, reveals a good agreement with the average of votes. For the region of small money we see that our model, like the data, predicts a constant behavior, $v \approx v_0$. After this region we have an evident correlation between votes and money. In order to better see the nuances of the correlated region we plotted in Figure 3(C) and (D) the normalized ratio $\langle v \rangle \Delta m/m$ for the eight largest Brazilian states. The symbols represents the data average and the lines show our model solution for each state identify by the color. For small and large values of m, we clearly see that our model correctly predicts the deviation from the linear behavior, $\langle v \rangle \Delta m/m \approx 1.0$.

3.5 Frequency distribution of votes

One of the first empirical investigations carried out by physicists concerning Brazilian elections was focused on the distribution $p(v)$ of the number of candidates receiving v votes. Since then, several other works have shown how the distribution changes for other countries [52, 59], and also proposed models to explain $p(v)$ [45, 40]. For our model the distribution of votes is an outcome of the distribution $p(m)$ of the number of candidates spending m. In Figure 4(A), (B) and (C) we show the distribution $p(m)$ for state deputies of São Paulo, Rio de Janeiro, and Minas Gerais, respectively. We see an apparent power law decay that expands over close of four orders of magnitudes. Using those distributions as input for our model we determine $p(v)$ for each one of those elections. In Figure 4(E), (F) and (G) we compare the empirical votes distribution for each state with the one obtained by our model. We see how our model reproduces correctly the empirical distribution for over two orders of magnitude. But even more, the correct prediction of $p(v)$ shows that the long tail, emblematic of $p(v)$, is not a consequence of networks or dynamic properties but mostly it comes from the money distribution.

3.6 Model validation

In order to test if our model really describe the data, we perform a series of comparisons using data from the 2014 state and federal deputy elections in the 26 states in Brazil. In Figure 5(A) each point is an election, and we show that our model calculates the turnout rate T in agreement with the data. This result was to be expected, since we

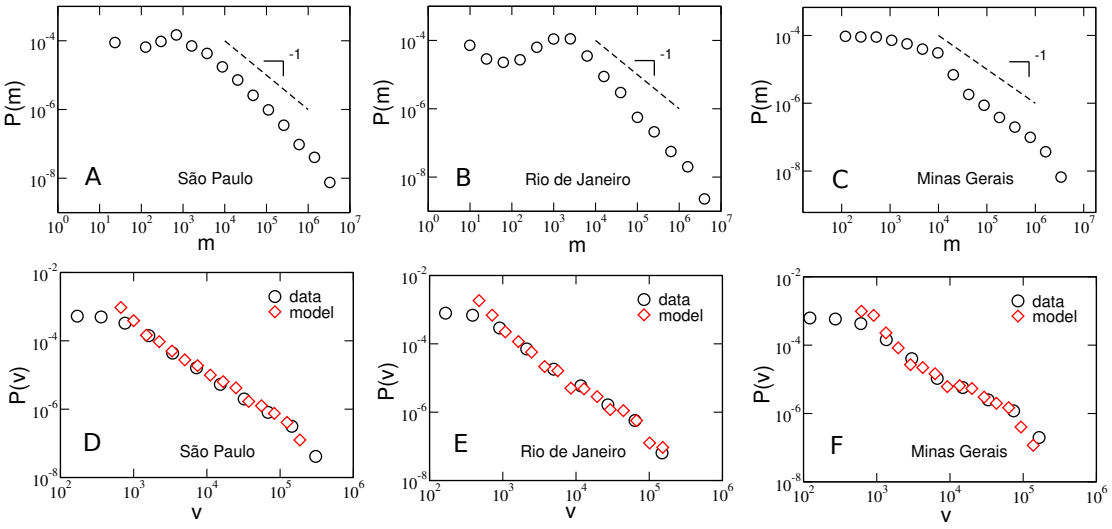

Figura 4: **Frequency distribution of money and number of votes.** In order to derive the distribution of votes our model takes as input the distribution of money. We see that the frequency distribution of money for the state deputies in São Paulo (A), Rio de Janeiro (B) and Minas Gerais (C) reveals a long tail characteristic that our model uses as an underlying cause for the observed vote distribution. We can now compare the actual distribution (black circles) of votes, $P(v)$, with the obtained by our model (red diamonds) for the election of state representatives in São Paulo (D), Rio de Janeiro (E), and Minas Gerais (F). We can see that our model have a good agreement with the data showing that the universal long tail characteristic of $p(v)$ is a direct consequence of money distribution and a competitive dynamic.

used Equation 3.3 to estimate Δm. We also compared how good our model estimates the candidate with the highest vote, v_{max}. Figure 5(B) shows the calculated v_{max} against the votes of the most voted candidate, where each black circles is an election. The red circles shows a linear approximation ($v_i \approx m_i/\Delta m$). We see that points of the mean-field model gather around the identity line showing a good agreement, demonstrating that our model is better than the linear model, which overestimate the votes (see section 3.8 for a statical comparison).

The Equation 3.3 shows that if we estimate Δm before the election, we can forecast the turnout ratio T. Fortunately, we can build some simple economical assumptions that presents a good agreement with the data. If we imagine that the candidates know the total money, M, and the size of the electorate, n, we can assume that the most simple and equitable division of votes will occur if $\Delta m = M/n$, in fact this is the first point where $T = 100\%$, which corresponds to the case without competition (more detail in section 3.7). Therefore we are assuming that the candidates adapt their strategy to have the maximum number of votes without have to compete and by consequence without have

to lose money. This heuristic gives us a value of valid votes of $T = 1 - e^{-1} \approx 63\%$. We show in Figure 5(C) the histogram of the number of total valid votes for the Congress elections in the years 2006, 2010 and 2014. The simple relation for T gives an impressive agreement with the average turnout, which we find experimentally to be $\approx 67\%$.

3.7 Study of the dispersion

. Our model allow us to calculate the mean or expected value of the number of votes. However, to fully describe the election we have also to study the statistical dispersion, which is given by the conditional probability distribution $p(v|m)$. We can use the concept of maximum entropy probability distribution (MaxEnt) in information theory to guess which is the $p(v|m)$ that maximize the Shannon's Entropy [16]. Imposing only a constraint for the mean $\langle v \rangle$ the maximum entropy continuous distribution is exponential,

$$p(v|m) = \frac{1}{\langle v \rangle} e^{-\frac{v}{\langle v \rangle}}. \tag{3.4}$$

The exponential distribution probability have the property that the mean and standard deviation are the same. We see in Figure 5(D) that our data show a tight relationship with approximately unit slope $\sigma \approx \langle v \rangle$, which strongly indicates that the Eq. 3.4 accounts for all the random variation on $v(m)$ with the expected value calculated by our model. In the inset of Figure 5(D), we show these two elements in a simulation for a election of state deputy for São Paulo's State. We can see that when we add the random dispersion our model have a remarkable resemblance with the data.

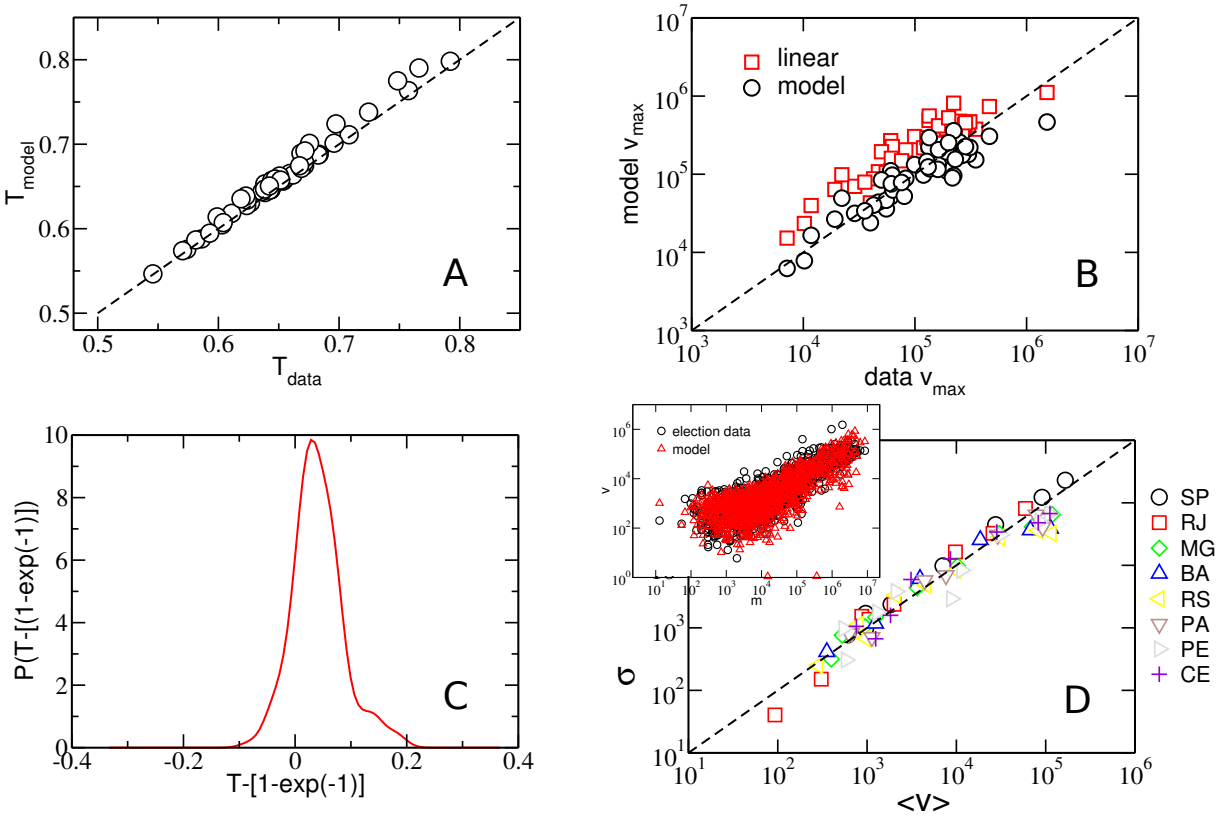

Figura 5: **Analytical results of the model.** By solving our model expressed in Eq. 1, we calculate the expected number of votes of each candidate. The total number of votes divided by the number of voters n is defined as the turnout ratio T. For all 56 parliamentary elections in 2014, we compared our model estimation of the turnout ratio, T_{model}, and the data ratio T_{data}, as we see in (A). The dashed line represents what would be the perfect agreement ($T_{data} = T_{model}$) and we see that the simulations (black circles) exhibit a good agreement. We can also select the candidate with the largest number of votes $v_{\max}$ and see how our model estimate this value. In (B) we see that our model (black circles) better estimates $v_{\max}$ than the linear (red squares), which most of the time overestimates it. We show in (C) a histogram of T for the election of 2006, 2010, and 2014. We find an average turnout value of $\approx 67\%$, which is consistent with our heuristic estimation of $T = 1 - e^{-1} \approx 63\%$. We know that the exponential distribution have the property that the mean and standard deviation are equal, this property can be used in order to test if the dispersion along the mean follows an exponential distribution. In (D) we see that for state deputies of the eight largest states in 2014 election the data is in close agreement with the hypotheses of $\sigma = \langle v \rangle$. We use the exponential distribution and the expected number of votes calculate by our model to generate a random election. We show in inset, for São Paulo, that when we add the random noise to our model (red triangle) we obtain a cloud that closely resembles the actual data (black circles).

3.8 Analytical Solution

Before solve the Equations 3.1 we can solve a simpler case, where we ignore the competition, analogous to an interaction in a physical system. If we suppose that the number of candidates N_C is much smaller than the number of voters n, we can neglect the interaction between candidates and the system of differential equations became uncoupled:

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = \frac{(n - v_i(t))}{n}[m_i(t) > 0] \tag{3.5}$$

These equations can be simply integrated, and the expectation for the number of votes in function of money is

$$v_i(m_i) = n - (n - v_i(0))e^{-m_i/(n\Delta m)}. \tag{3.6}$$

We can assume that $n\Delta m >> m_i$ for any i, which means that the candidate with the largest campaign finance will not have enough money to reach out the whole system. Therefore, we can find a linear approximation for the dependence between votes and money:

$$v_i(m_i) = n - (n - v_i(0))\left(1 - \frac{m_i}{n\Delta m} + \frac{m_i^2}{2\Delta m^2 n^2} - ...\right), \tag{3.7}$$

$$v_i(m_i) \approx v_i(0) + (n - v_i(0))\frac{m_i}{n\Delta m}. \tag{3.8}$$

Assuming $n - v_i(0) \approx n$, we finally find the linear approximation

$$v_i(m_i) \approx v_i(0) + m_i/\Delta m. \tag{3.9}$$

Now we want to solve the competition equation system. Taking Equation 3.1 and summing over i we can find a differential equation for $S(t)$ that can be easily integrated:

$$\sum_{i=0}^{N_c} \frac{\mathrm{d}v_i}{\mathrm{d}t} = \frac{\mathrm{d}S}{\mathrm{d}t} = \frac{(n - S(t))}{n} \sum_{i=0}^{N_c}[m_i(t) > 0] \tag{3.10}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{(n - S(t))}{n} r(t) \tag{3.11}$$

$$S(t) = n - (n - S(0)) \exp\left(-\frac{1}{n}\int_0^t r(t')dt'\right). \tag{3.12}$$

Where $r(t) = \sum_i [m_i(t) > 0]$ and can be interpreted as the number of candidates who still have money at time t, which depends solely on the distribution of money.

The Equation 3.12 enables us to determine the turnout rate T of the election in function of Δm, the total money $M = \sum_i m_i$, and n. We just need to take the limit $t \to \infty$ and by doing that we determine the value where S saturates.

$$T = \frac{S(t \to \infty)}{n} = 1 - e^{-M/(n*\Delta m)}. \tag{3.13}$$

To find Equation 3.13 we used the fact that $\int_0^\infty r(t')dt' = M/\Delta m$ and we use the approximation $n - S(0) \approx n$. In order to understand this integral we have to remember the definition of $r(t)$ and also the fact that $dt = -dm/\Delta m$. So the integral became

$$\int_0^\infty r(t')dt' = \int_0^\infty \sum_{i=0}^{N_c} [m_i(m') > 0] dm'/\Delta m, \tag{3.14}$$

$$\int_0^\infty r(t')dt' = \sum_{i=0}^{N_c} m_i/\Delta m = M/\Delta m. \tag{3.15}$$

In Figure 6(A) we show the turnout rate T (Eq. 3.13) in function of Δm, and compare with the case without competition (T_{linear}), which is obtained by simple summing Equation 3.9 and taking $S(0) \approx 0$. The amount of votes (or money) lost by competition can be evaluated by looking at the subtraction of the two rates. We see that there is a maximum loss when $\Delta m = M/n$. As we can observe, Δm can work physically as an external field that will provide energy to the system in order to convince a fraction of the voters, or also Δm can be seen as a Temperature, wherein for small values the system organize in a big cluster of convinced nodes and for large values of Δm the system stay in a "disorganized" state ($T \approx 0$).

Now we can integrate our model in function of $S(t)$ and calculate v,

$$v_i = v_i(0) + \frac{m_i}{\Delta m} - \frac{1}{\Delta m}\int_0^{m_i} \frac{S(m')}{n} dm', \tag{3.16}$$

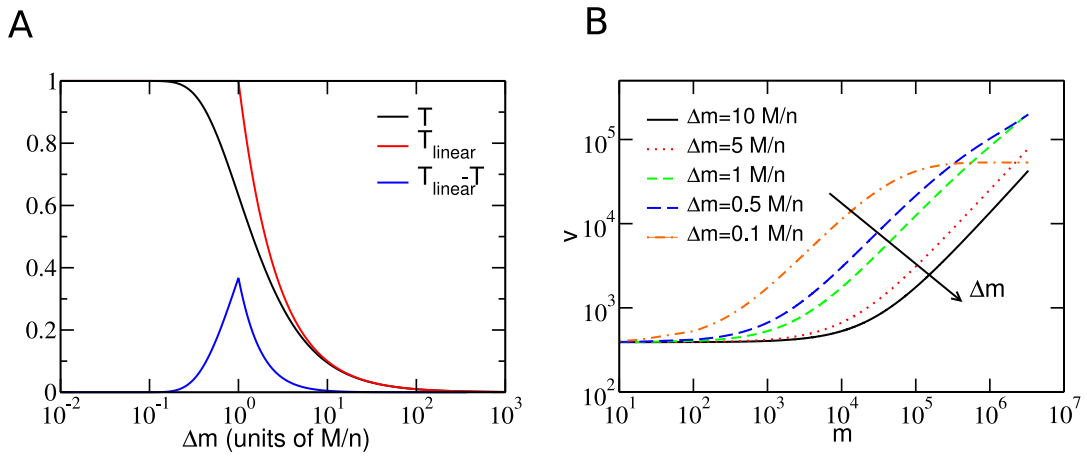

Figura 6: **Dependence with Δm.** The solution of the mean field model enables us to calculate the turnout radio T in function of the adimensional $n\Delta m/M$ parameter. In (A) we compare turnout for the approximated case where we excluded the competition between the candidates, T_{gas}, with the case with competition, T. The competition creates an exponential saturation, which increases the loss of money when candidates seek new voters. This inefficiency is maximum when $n\Delta m/M = 1.0$, we can see that by looking for $T_{gas} - T$. We can also vary Δm and see how the curve $v(m)$ change. In (B) we show that as we decrease Δm the values of $v(m)$ usually increases, as expected by the definition of Δm, however there is a point that we have a saturation process as the total number of votes starts to get close of the size of the system ($T \to 1.$).

$$v_i(m_i) = v_i(0) + \frac{1}{\Delta m}\left(1 - \frac{S(0)}{n}\right)\int_0^{m_i} \exp\left[-\frac{1}{n\Delta m}\int_0^{m'} r(m'')dm''\right] dm'. \tag{3.17}$$

To find an analytical expression for v we have to solve the integral $\int_0^{m'} r(m'')dm''$ inside the exponential. But first we can decompose the external integral as

$$\begin{aligned} v_i = v_i(0) &+ \frac{1}{\Delta m}\left(1 - \frac{S(0)}{n}\right)\int_0^{m_{i-1}} \exp\left[-\frac{1}{n\Delta m}\int_0^{m'} r(m'')dm''\right] dm' \\ &+ \frac{1}{\Delta m}\left(1 - \frac{S(0)}{n}\right)\int_{m_{i-1}}^{m_i} \exp\left[-\frac{1}{n\Delta m}\int_0^{m'} r(m'')dm''\right] dm', \end{aligned} \tag{3.18}$$

$$v_i = v_i(0) - v_{i-1}(0) + v_{i-1} + \frac{1}{\Delta m}\left(1 - \frac{S(0)}{n}\right)\int_{m_{i-1}}^{m_i} \exp\left[-\frac{1}{n\Delta m}\int_0^{m'} r(m'')dm''\right] dm'. \tag{3.19}$$

The result of this integral rely on the limits of the external integral. Using the

definition of $r(m)$ for the external interval $m' \in [m_{i-1}, m_i]$ we have that

$$\int_0^{m'} r(m'')dm'' = m_0 + m_1 + m_2 + ... + m_{i-1} + (N_c - i)m'. \quad (3.20)$$

By solving the integrals we finally find that the number of votes v_i is given by

$$v_i = v_i(0) - v_{i-1}(0) + v_{i-1} - \frac{n - S(0)}{N_c - i} e^{-\sum_{j=0}^{i-1} m_j/(n\Delta m)} [e^{-\frac{(N_c-i)m_i}{n\Delta m}} - e^{-\frac{(N_c-i)m_{i-1}}{n\Delta m}}]. \quad (3.21)$$

We can see in Equation 3.21 that the number of votes of a candidate i is not only a function of his budget m_i but indeed it depends of the set $\{m_i\}$. We show in Fig. 6(B) how the curves $v(m)$ changes in function of Δm. As we decrease Δm a large fraction of the system gets occupied, $T \to 1.0$, and $v(m)$ shows a curvature for larges values of m and start to saturate displaying a nonlinear growth.

3.9 Comparing models

In order to compare the model with the simple case without competition we used the Akaike's Information Criterion (AIC) [63]. AIC uses information theory to gives a relative estimation of the information lost when a given model is used to represent the process that generates the data. We used AIC to measure the relative quality of our model when we compare with the linear non-competitive model.

Lets suppose that we have a model, with k parameters, that fits a data set with N points. If S is the residual sums of squares then the AIC is defined by

$$AIC = N \ln\left(\frac{S}{N}\right) + 2(k+1). \quad (3.22)$$

Using the Equation 3.22 we can calculate the AIC for each model and the preferred model is the one with the minimum AIC value. We define model A as the model without competition and B the most complex model, where there is competition. The difference in AIC is defined as $\Delta AIC = AIC_B - AIC_A$. With this difference we can calculate the probability that model B minimize the information loss:

$$P_B = \frac{e^{-0.5\Delta AIC}}{1 - e^{-0.5\Delta AIC}}. \quad (3.23)$$

As well as, the probability of model A, $P_A = 1 - P_B$. We can also divide the probability for model B by the probability for the model A to obtain the evidence radio, which means how many times is B more likely to minimize the information loss.

We performed all these analysis for Federal and State deputies in 26 Brazilian states. The data was compared with the model without any transformation and also after application of logarithm. The results for the 2014 elections can be seen in Tables 1, 2, 3, and 4. The AIC shows that the model with competition best explains the data when compared to a linear model for all cases.

Tabela 1: We used the Akaike's information criterion (AIC) to compare the two models: A (without competition) and B (with competition). The AIC lets us determine which model is more likely to describe correctly the data and quantify by calculating the probabilities and an evidence radio. The probability column shows the likelihood of each model to be the most correctly. The evidence radio is the fraction of Probability B by Probability A, which means how many times model B is likely to be correct than model A. The AIC was applied in the log(data).

Federal Deputies				
State	ΔAIC	Probability A	Probability B	Evidence radio
AC	6.1827384262	0.0434646736	0.9565353264	22.0071898888
AL	4.7608349884	0.0846782011	0.9153217989	10.8094147898
AM	2.6647303199	0.2087684091	0.7912315909	3.7899967439
AP	2.806065353	0.1973353125	0.8026646875	4.0675167435
CE	10.5123485714	0.0051881611	0.9948118389	191.7465189
ES	14.4468382526	0.0007287731	0.9992712269	1371.1692559
GO	7.4677125018	0.0233425922	0.9766574078	41.8401435584
MA	7.7592626578	0.0202403068	0.9797596932	48.4063657539
MS	10.1109475602	0.0063339711	0.9936660289	156.878838813
MT	3.9037010608	0.1243517168	0.8756482832	7.0417064229
PA	9.663695483	0.0079087312	0.9920912688	125.442532017
PB	4.6657768971	0.0884355349	0.9115644651	10.3076717549
PI	2.3415470513	0.2367151943	0.7632848057	3.2244858966
RN	5.1455026334	0.0709128217	0.9290871783	13.1018221599
RO	9.807097669	0.0073655493	0.9926344507	134.767198525
RR	1.8879845033	0.2800946322	0.7199053678	2.5702219362
SC	13.3469679287	0.0012623917	0.9987376083	791.147136976
SE	2.8635043257	0.1928258238	0.8071741762	4.1860273715
TO	5.8040508651	0.0520535295	0.9479464705	18.2109931789
BA	16.0404485774	0.0003286382	0.9996713618	3041.85951117
MG	32.2756994687	9.80439653939e-08	0.999999902	10199504.8644
SP	74.5043113536	6.63123395728e-17	1	1.50801495837e+16
RJ	42.5533963117	5.74972928891e-10	0.9999999994	1739212316.23
RS	18.8727293265	7.97635097082e-05	0.9999202365	12536.0611657
PE	7.192496085	0.026694303	0.973305697	36.4611767018
PR	15.899933541	0.0003525495	0.9996474505	2835.48072726

Tabela 2: We used the Akaike's information criterion (AIC) to compare the two models: A (without competition) and B (with competition). The AIC lets us determine which model is more likely to describe correctly the data and quantify by calculating the probabilities and an evidence radio. The probability column shows the likelihood of each model to be the most correctly. The evidence radio is the fraction of Probability B by Probability A, which means how many times model B is likely to be correct than model A. The AIC was applied in the log(data).

States Deputies				
State	ΔAIC	Probability A	Probability B	Evidence radio
AC	40.1333350824	1.92822192099e-09	0.9999999981	518612503.667
AL	10.5644673228	0.005055382	0.994944618	196.808989398
AM	25.7812709128	2.52154694444e-06	0.9999974785	396580.948318
AP	13.4280658125	0.0012122878	0.9987877122	823.886605911
CE	24.7191563734	4.28846166679e-06	0.9999957115	233182.849525
ES	40.567665947	1.55182686878e-09	0.9999999984	644401781.26
GO	22.588786324	1.24423372754e-05	0.9999875577	80369.751722
MA	17.3075446836	0.000174437	0.999825563	5731.72803942
MS	29.9037339555	3.20986330536e-07	0.999999679	3115396.46359
MT	19.5744050013	5.61626470736e-05	0.9999438374	17804.4285563
PA	31.4296195953	1.49673445309e-07	0.9999998503	6681210.87384
PB	18.9409648861	7.7088261887e-05	0.9999229117	12971.1435601
PI	8.8288917589	0.0119565683	0.9880434317	82.6360378881
RN	8.8191591598	0.0120141936	0.9879858064	82.2348830361
RO	24.1600601847	5.67162001552e-06	0.9999943284	176315.466418
RR	20.058166039	4.4096633465e-05	0.9999559034	22676.468129
SC	30.7721830943	2.07924302228e-07	0.9999997921	4809441.61582
SE	8.7956277831	0.0121546554	0.9878453446	81.2730027198
TO	13.5362503146	0.0011485278	0.9988514722	869.679851625
BA	42.8429750252	4.97469178945e-10	0.9999999995	2010174784.34
MG	55.735515674	7.89199040127e-13	1	1.26710747119e+12
SP	168.97837878	2.0268017352e-37	1	4.93388170453e+36
RJ	82.3116713751	1.33735794708e-18	1	7.47742967531e+17
RS	52.8467317602	3.3456307332e-12	1	298897302106
PE	19.5248749346	5.75708013522e-05	0.9999424292	17368.9162859
PR	42.819329153	5.0338563126e-10	0.9999999995	1986548557.2

Tabela 3: We used the Akaike's information criterion (AIC) to compare the two models: A (without competition) and B (with competition). The AIC lets us determine which model is more likely to describe correctly the data and quantify by calculating the probabilities and an evidence radio. The probability column shows the likelihood of each model to be the most correctly. The evidence radio is the fraction of Probability B by Probability A, which means how many times model B is likely to be correct than model A.

Federal Deputies

State	ΔAIC	Probability A	Probability B	Evidence radio
AC	50.1494383861	1.28880680099e-11	1	77591148589.4
AL	101.196392792	1.0604312338e-22	1	9.43012585939e+21
AM	107.771152638	3.96087877268e-24	1	2.524692265e+23
AP	120.857209596	5.70414277247e-27	1	1.75311179942e+26
CE	184.905492905	7.05151410541e-41	1	1.41813514807e+40
ES	202.215834338	1.22854055302e-44	1	8.13973944563e+43
GO	79.8068548356	4.67909285203e-18	1	2.13716639448e+17
MA	224.617775801	1.67830046988e-49	1	5.95840862792e+48
MS	118.603128883	1.76058823276e-26	1	5.67991982108e+25
MT	120.825144982	5.79633035462e-27	1	1.72522947938e+26
PA	141.83521874	1.58808440446e-31	1	6.29689453025e+30
PB	125.232526939	6.39885542365e-28	1	1.56277948757e+27
PI	105.735798465	1.09588023355e-23	1	9.12508474362e+22
RN	103.807827371	2.87353636201e-23	1	3.48003252446e+22
RO	92.3868106145	8.6787858999e-21	1	1.1522348996e+20
RR	83.7990976172	6.35707240879e-19	1	1.57305114005e+18
SC	130.020826581	5.83896996879e-29	1	1.71263083274e+28
SE	72.7348067195	1.60633972519e-16	1	6.22533318649e+15
TO	55.1303103352	1.06808352921e-12	1	936256362587
BA	342.184328822	4.96154688314e-75	1	2.01550045491e+74
MG	777.261043094	1.65923917468e-169	1	6.02685866668e+168
SP	749.737824971	1.57217132373e-163	1	6.36062994474e+162
RJ	901.70236979	1.57695115134e-196	1	6.34135051776e+195
RS	451.109474546	1.10362679252e-98	1	9.06103409936e+97
PE	142.356671451	1.22360590247e-31	1	8.1725660033e+30
PR	334.950626527	1.84669679188e-73	1	5.41507411718e+72

Tabela 4: We used the Akaike's information criterion (AIC) to compare the two models: A (without competition) and B (with competition). The AIC lets us determine which model is more likely to describe correctly the data and quantify by calculating the probabilities and an evidence radio. The probability column shows the likelihood of each model to be the most correctly. The evidence radio is the fraction of Probability B by Probability A, which means how many times model B is likely to be correct than model A.

States Deputies

State	ΔAIC	Probability A	Probability B	Evidence radio
AC	576.061906458	8.12356005108e-126	1	1.23098739187e+125
AL	238.628928458	1.52190160204e-52	1	6.57072703427e+51
AM	682.650418552	5.81226058412e-149	1	1.72050097467e+148
AP	358.738756255	1.26144655989e-78	1	7.92740677087e+77
CE	420.054263752	6.11470593755e-92	1	1.63540162064e+91
ES	480.640448515	4.26827816914e-105	1	2.34286510947e+104
GO	989.587809594	1.29938385781e-215	1	7.69595523285e+214
MA	519.730902297	1.38633608904e-113	1	7.21325808299e+112
MS	439.886900022	3.01837593293e-96	1	3.31303993347e+95
MT	310.209118004	4.35457632874e-68	1	2.29643465749e+67
PA	685.433108646	1.44574467234e-149	1	6.9168506662e+148
PB	400.461973128	1.09846804884e-87	1	9.10358750133e+86
PI	189.012621263	9.0454627114e-42	1	1.10552664016e+41
RN	249.978331952	5.2226980642e-55	1	1.91471915035e+54
RO	482.146906385	2.00969219136e-105	1	4.97588637852e+104
RR	370.038828903	4.4369982636e-81	1	2.25377595525e+80
SC	462.924309949	3.00098154818e-101	1	3.3224308096e+100
SE	139.093088068	6.2563306112e-31	1	1.59838100341e+30
TO	282.919902956	3.67048676832e-62	1	2.72443428656e+61
BA	642.920743716	2.46339667887e-140	1	4.0594355289e+139
MG	1450.44716375	1.09496501832e-315	1	inf
SP	2129.42600533	0	1	inf
RJ	1694.58149782	0	1	inf
RS	618.918508849	4.01377875167e-135	1	2.49141784306e+134
PE	369.31393498	6.37526107172e-81	1	1.56856321451e+80
PR	784.782590509	3.8603414867e-171	1	2.59044440354e+170

3.10 Discussion

The nonlinear relation between v and m in real elections as a consequence of competition between candidates can complement other statistical analyses for political campaign and electoral outcome [55, 54, 60]. These analyses detect a number of statistical patterns of electoral processes, such as the relations between party size and temporal correlations [61], the relations between the number of candidates and voters [58], and the distribution of votes [44, 56, 40, 52, 59]. Our analysis goes beyond these approaches by providing possible answers for a number of key issues: a simple equation that rationalize statistical evidence about the effects of political campaign on electoral outcome, an estimation for the expected number of votes as a function of campaign money by accurately predicting the vote distribution, and finally a simple heuristic that can estimate the electorate turnout rate.

A close inspection of the campaign data reveals a ubiquitous nontrivial relation between v and m for all elections investigated. We showed that this relation is an unambiguous sublinear correlation between the money spent by candidate and her/his number of votes v, specially for the the top spender candidates, revealing that the electoral process works in a diseconomy of scale state.

To explain the diseconomy of scale in the campaign economy, we proposed a mean field model where candidates compete with each other and must spend their money in order to get votes. Despite its simplicity, the model proved capable of reproducing the complexity of the dependence of v with respect of m. This good agreement makes our model a possible alternative to study other aspects of human collective behaviour involving, for example, diffusion of innovation and decision-making, such as the competition in market share where companies invest in advertising for products [62].

4 STATISTICAL SIGNS OF SOCIAL INFLUENCE ON SUICIDES

4.1 Introduction

It is not uncommon in nature to observe properties that present non-trivial forms of scale dependence. This is the case, for instance, of critical phenomena, where scaling invariance, universal properties and renormalization concepts constitute the theoretical framework of a well-established field in physics [69]. In biology, the so-called allometric relations certainly represent outstanding examples of natural scaling laws. Precisely, allometry implies the use of power-laws, $Y \propto M^{\beta}$, to describe the dependence of a wide range of anatomical, physiological and behavioral properties, denoted here as Y, on the size or the body mass of different animal species, M. If the scaling exponent is $\beta = 1$, the variables Y and M are trivially proportional, and the relation between them is said to be isometric, while $\beta \neq 1$ indicates an allometric type of relationship. The so-called "three-quarters law" or Kleiber's law, as originally proposed by the agricultural biologist Max Kleiber in 1947 [5], is surely one of the most prominent allometric relations found in natural sciences. Based on an extensive set of experimental data, this fundamental law states that the metabolic rate of all animals should scale to the 3/4 power of their corresponding masses [70, 4].

In analogy with biological scaling laws, Bettencourt *et al.* [6] showed that, regardless the enormous complexity and diversity of human behavior and the extraordinary geographic variability of urban settlements, cities belonging to the same urban system obey pervasive allometric relations with population size, therefore exhibiting nonextensive rates of innovation, wealth creation, patterns of consumption, human social behavior, and several other properties related to the urban infrastructure. The authors then conclude that all data can be grouped into three categories, namely, material infrastructure, individ-

ual human needs, and patterns of social activity. Despite the unambiguous presence of power-laws, the urban indicators do not necessarily follow a universal behavior. Instead the results can be divided in three different classes. The isometric (linear) case ($\beta = 1$) typically reflects the scaling of individual human needs, like the number of jobs, houses, and water consumption. As in biology, the allometric sublinear behavior ($\beta < 1$) implies an economy of scale in the quantity of interest, because its *per capita* measurement decreases with population size. In the case of cities, this is materialized, for example, in the number of gasoline stations, the total length of electrical cables, and the road surface (material and infrastructure). The case of superlinear allometry ($\beta > 1$) in urban indicators emerges whenever the complex patterns of social activity have significant influence. Wages, income, growth domestic product, bank deposits, as well as rates of invention, measured by new patents and employment in creative sectors, all display a superlinear increase with population size [6]. While these results indicate that larger cities are associated with higher levels of human productivity and quality of life, superlinear scaling can also characterize negative urban scenarios, such as the prospect of living costs, crime rates, pollution and disease incidence [6, 71, 72, 68, 73].

4.2 Materials and Methods

Brazilian Data. We analyzed data available for all Brazilian cities from 1992 to 2009, made freely available by the Brazil's public healthcare system – DATASUS [74]. Here we consider that cities are the smallest administrative units with local government. The data consist of the number of homicides, suicides, and deaths in traffic accidents as well as the population for each city.

US Data. We used data from the National Cancer Institute SEER, Surveillance Epidemiology and End Results downloaded from http://seer.cancer.gov/data/. The Institute provides mortality data aggregated for three or five years. To compare with Brazil, we use suicide mortality data for each American county and MSA (Metropolitan Statistical Area) accumulated through the five years of 2003 to 2007. Since the population is almost constant for a five years period, we adopted the average population as a measurement for the population on the allometry relation.

Exponent Determination. In order to reduce fluctuations, we first apply a non-parametric fitting method to the data, namely, the Nadaraya-Watson kernel regression [79, 80]. We also compute the 95% ($\alpha = 0.05$) confidence interval (CI) by the so-called

$\alpha/2$ quantile function over 300 random bootstrapping samples. The power-law exponent is calculated by performing an ordinary least square (OLS) fitting [81] over the results of the Nadaraya-Watson kernel regression in the population interval $[10^4, 10^6]$ for Brazilian cities, and $[10^5, 10^7]$ for American MSAs.

4.3 Allometry in Urban Indicators

The main goal here is to investigate the scaling behavior with city population of three urban indicators, namely the number of homicides, deaths in traffic accidents and suicides. For this, we analyzed data available for all Brazilian cities and as well as suicide for US counties and MSAs, as presented previously on Materials and Methods. For 2009 in Brazil, as shown in Fig. 7, the increase in the number of casualties D with city population P for the three death causes can be properly described in terms of power-laws, $D = D_0 P^{\beta}$, where D_0 is a normalization constant, and the exponent β reflects a global property at play across the urban system. Interestingly, while the number of deaths by traffic accidents display isometric scaling, $\beta \approx 1$, homicides and suicides are both allometric, but obeying superlinear and sublinear scaling with population, respectively. These results suggest that the decision to commit a crime or to suicide, instead of being purely a consequence of individual choices, might have strong correlations with the underlying complex social organization and interactions. This does not seem to be the case of traffic accidents, since the strong evidence for isometric scaling, $\beta = 0.99 \pm 0.02$, indicates that such events should result from random processes, i.e., no social relations need to be implied among the involved people.

From the superlinear scaling exponent found for the number of homicides in Brazil at the year 2009, $\beta = 1.24 \pm 0.01$, one should expect that, by doubling the population of a city, the number of homicides would grow approximately by a factor of 135% in average, instead of just growing 100%, as if we had an isometric scaling. This super-linear behavior is consistent with that found for serious crimes in USA [6]. The result obtained for suicide scaling seems quite surprising. As depicted in Fig. 7C, the number of suicides scales with the population size as a power law with exponent $\beta = 0.84 \pm 0.02$, which implies an "economy of scale" of 22% in comparison with isometric scaling, similar to Kleiber's law for metabolic rates and animal masses. This sublinear behaviour is reminiscent of the seminal study by Émile Durheim [67], one of the fathers of modern sociology. In his book *Suicide*, Durkheim explored the differences in average suicide rates among Protestants and Catholics, arguing that stronger social control among Catholics leads to lower suicide

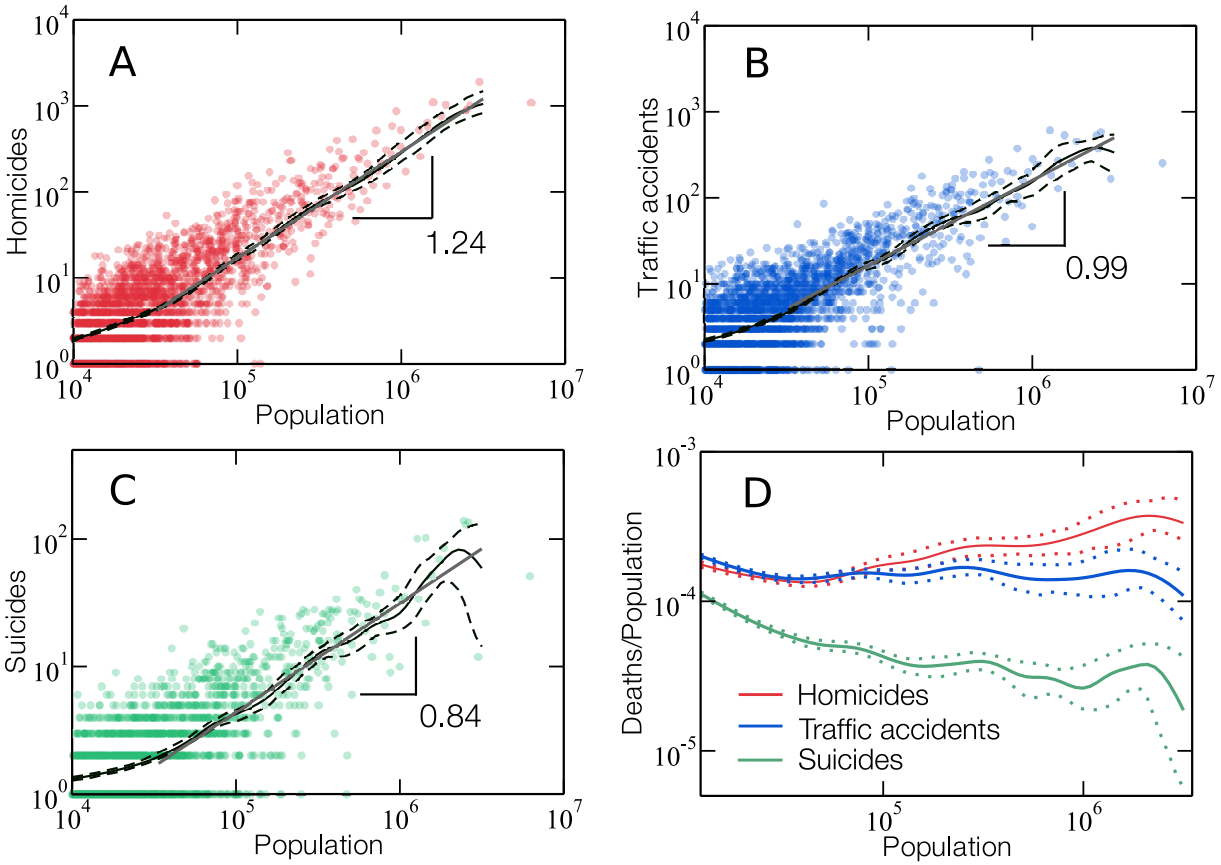

Figura 7: **Scaling relations for homicides, traffic accidents, and suicides for the year of 2009 in Brazil.** The small circles show the total number of deaths by (A) homicides (red), (B) traffic accidents (blue), and (C) suicides (green) vs the population of each city. Each graph represents only one urban indicator, and the solid gray line indicate the best fit for a power-law relation, using OLS regression, between the average total number of deaths and the city size (population). To reduce the fluctuations we also performed a Nadaraya-Watson kernel regression [79, 80]. The dashed lines show the 95% confidence band for the Nadaraya-Watson kernel regression. The ordinary least-squares (OLS) [81] fit to the Nadaraya-Watson kernel regression applied to the data on homicides in (A) reveals an allometric exponent $\beta = 1.24 \pm 0.01$, with a 95% confidence interval estimated by bootstrap. This is compatible with previous results obtained for U.S. [6] that also indicate a super-linear scaling relation with population and an exponent $\beta = 1.16$. Using the same procedure, we find $\beta = 0.99 \pm 0.02$ and 0.84 ± 0.02 for the numbers of deaths in traffic accidents (B) and suicides (C), respectively. This non-linear behavior observed for homicides and suicides certainly reflects the complexity of human social relations and strongly suggests that the the topology of the social network plays an important role on the rate of these events. (D) The solid lines show the Nadaraya-Watson kernel regression rate of deaths (total number of deaths divided by the population of a city) for each urban indicator, namely, homicides (red), traffic accidents (blue), and suicides (green). The dashed lines represent the 95% confidence bands. While the rate of fatal traffic accidents remains approximately invariant, the rate of homicides systematically increases, and the rate of suicides decreases with population.

rates. The crucial contribution from Durkheim was certainly to treat the suicide as a social fact, by explaining variations in its rate at a macro level as a direct consequence of society-scale phenomena, such as lack of connections between people (group attachment) and lack of regulations of behavior, rather than individuals' feelings and motivations.

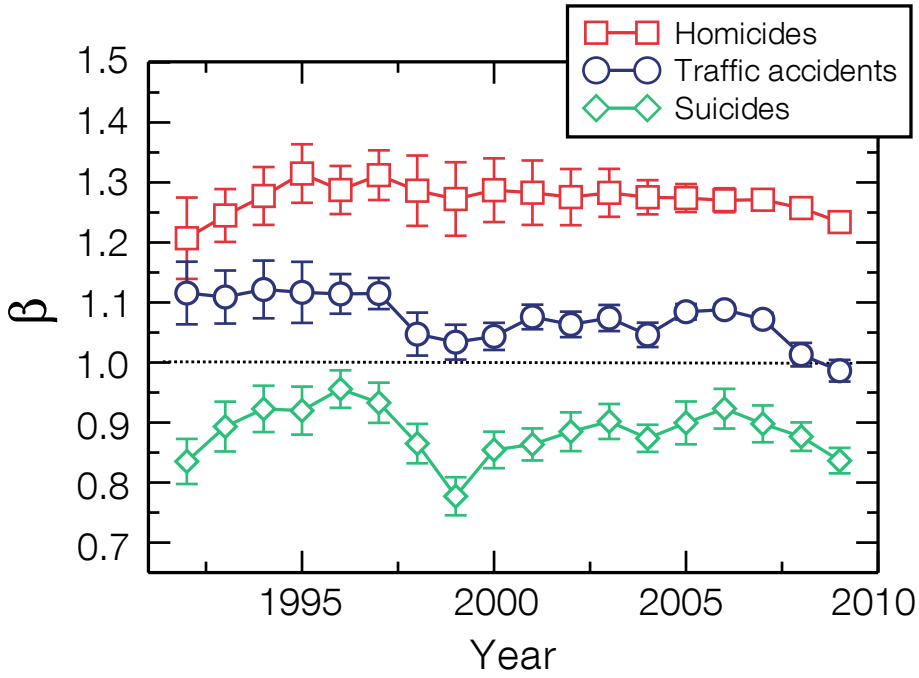

Figura 8: **Temporal evolution of allometric exponent β for homicides (red squares), deaths in traffic accidents (blue circles), and suicides (green diamonds).** Time evolution of the power-law exponent β for each behavioral urban indicator in Brazil from 1992 to 2009. We can see that the non-linear behavior for homicides and suicides are robust for this 19 years period, and for the traffic accidents the exponent remain close of 1.0.

The discrepancies observed in the scaling behaviors of homicides, deaths in traffic accidents and suicides become even more evident if we plot the average number of deaths per capita against city population, as shown in Fig. 7D. Under this framework, the systematic decrease in suicide rate with population indicates that a large supply of potential social contacts and interactions might work as an "antidote" for this tragic event. This result is consistent with the idea that human happiness is more a collective phenomenon than a consequence of individual well-being conditions. In analogy with health, it is then possible to consider a "happiness epidemy" spreading in a social network like showed previously in [75]. In Fig. 8 we show the dependence on time of the exponent β for a period of 18 years, from 1992 to 2009. We see a robust behavior for β, in such a way that even for different years we are still having $\beta > 1.0$ for homicides, $\beta < 1.0$ for suicides, and β slightly above 1.0 for deaths by traffic accidents.

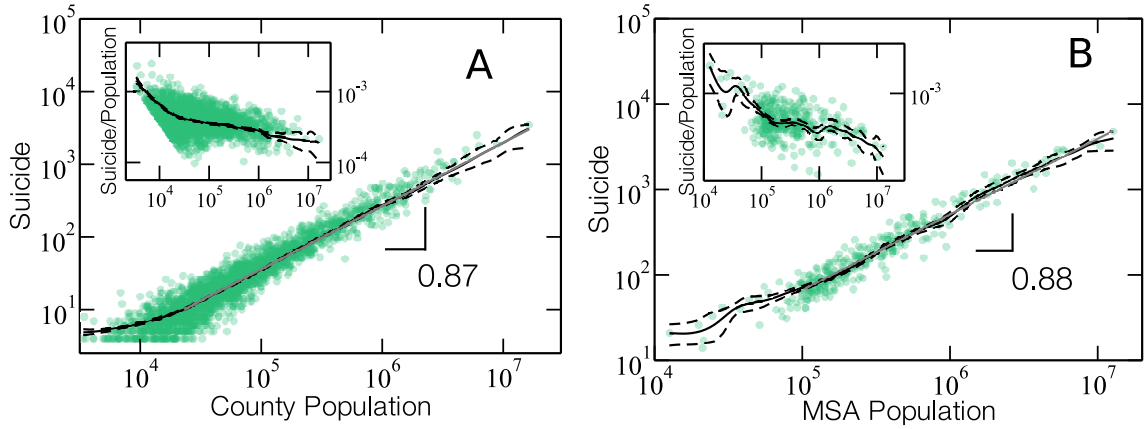

Figura 9: **Scaling relationship between suicides and population for US counties and MSAs.** The small circles show the total number of suicides over five years (2003 to 2007) vs the average population for counties (A) and MSAs (B). The solid gray line indicate the best fit of a power law, using OLS regression, between the average total number of suicides and population. The dashed black lines delimit the 95% confidence band given by the Nadaraya-Watson kernel regression (solid black line) [79, 80]. The allometric exponents are obtained through an ordinary least-squares (OLS) fit [81] over the Nadaraya-Watson kernel regression applied to the suicides data. We find $\beta = 0.87 \pm 0.01$ for counties and $\beta = 0.88 \pm 0.01$ for MSAs with a 95% confidence interval estimated by bootstrap. The insets in each graph show the systematic decreases of suicide rates with population in both cases.

We also analyzed data available for suicides in all US counties and MSAs. The data is an accumulation of the total number of suicides during a period of five years, from 2003 to 2007. In Figs. 9(A) and 9(B) we show the dependence of the total number of suicides during these five years on the average population for each county and MSA, respectively. As depicted, the number of suicides also scales with a sub-linear power law with exponent $\beta = 0.87 \pm 0.01$ for counties, and $\beta = 0.88 \pm 0.01$ for MSAs, which are in agreement with our previous results for Brazil.

In Fig. 10 we show the density map in the 2D plane, fatality per capita (deaths divided by population) versus population size for Brazil in 2009. To obtain the approximate density we perform kernel density estimation in the log-log space. We also include in the figure lines indicating the 10%, 50%, and 90% levels for each population size. Besides confirming the superlinear, linear, and sublinear behaviors, this results also show how the probability distribution of rates of fatality vary with the population size. Also, the 10% and 90% lines are representative of expected extreme cases of low and high fatalities, respectively.

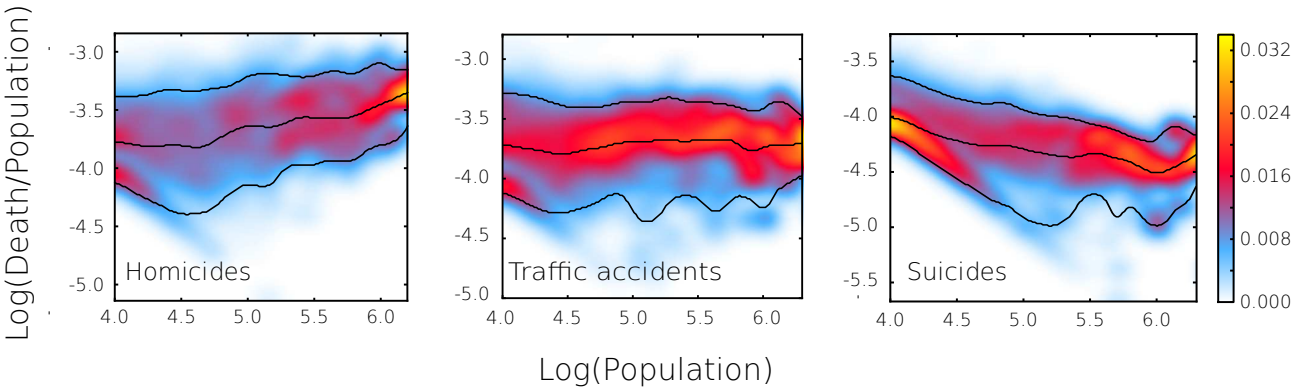

Figura 10: **Fatality per capita versus population for homicides, traffic accidents, and suicides.** The color map represents the conditional probability density obtained by kernel density estimation. The bottom and top lines correspond to the 10% and 90% bounds of the distribution for each population size, that is 80% of the sampled points are between these lines. The middle line is the 50% level or "median"expected for each population size. The diagonal shape observed in the left side of density maps are cases of low number of fatal events, one or two fatalities. After this region we observe that the three level lines wiggle around an average power-law behavior. In the case of homicides the three level lines indicate an increase in the expected density of fatality with the population size. Similarly, for traffic accidents the lines are close to horizontal, that is, the probability distribution for the rate of fatality is near independent of the population. For suicides, the median show a slight decrease with population size, while the 90% level, that is associated with cases of extreme rates of suicide, show a pronounced decrease. The sublinear growth observed for suicides, as depicted in Fig. 7C, is likely due to the suppression of these extremely high rates in large urban areas.

4.4 Discussion

Allometric relations are ubiquitous in Nature, appearing in a wide variety of biological, sociological, chemical and physical systems [76, 6, 69]. Even the arrangement of Lego pieces has been recently reported to obey a sublinear scaling between the number of pieces types vs. the total number of pieces for many Lego sets [77]. As originally proposed in biology by Kleiber [5] and extended later by others [76], a single power law comprising 27 orders of magnitude can associate the metabolisms of a microscopical bacteria and a blue whale, weighting a few picograms and more than 100 t, respectively. Unlike allometric relations in Biology, few attempts have been made to explain the origin of universal scaling laws describing urban indicators. In a recent study [25], a quantitative framework has been developed to consider the interactions between people over a social network that is capable to predict the allometric scaling for urban systems.

Here we studied the scaling relation with population of three behavioral urban indi-

cators, namely, number of homicides, victims of vehicle crashes, and suicide events. We show that, unlike the incidence of vehicle crashes with fatal victims, which exhibits isometric (linear) scaling, homicides and suicides are characterized by allometric behaviors, sublinear and superlinear, respectively.

Precisely, we found that the superlinear scaling relation for the number of homicides has an allometric exponent that lies between 1.21 and 1.31 along the years, between 1992 and 2009. Despite all their positive attributes [71, 72] (e.g., higher incomes and levels of creative activities), these results show that larger and urbanized cities also have a dark side, namely, higher levels of violence. Interestingly, the effect is exactly the opposite in the number of suicide events, which typically follow a sublinear scaling with population, with an allometric exponent value that varies in the range between 0.78 and 0.95. Such a result led us to the conclusion that a suicide event should not be taken as an isolated individual decision. This is consistent with the conceptual framework put forward by Durkheim [67], under which suicides need to be treated as social facts, actually affected by complex human relations. The sublinear behavior found can be even considered as a counter-intuitive result. The more common and straightforward view would be to associate the suicide causes uniquely to a health condition of psychological or mental illness that could nevertheless be strongly linked to external urban factors. For instance, traffic jams, pollution, and stressful jobs all create a harmful environment for the population. However, the fact that we found sublinearity, namely, that the number of suicides is disproportionately small for larger cities, discloses an entirely different perspective. We conjecture that this property can be intimately related with an "emotional epidemy" as previously hypothesized in Ref. [75, 78]. This phenomenon can explain the systematic attenuation through the social network of contagious emotions states of potentially suicidal individuals.

5 CONCLUSIONS

In this thesis, we apply a "physicist" approach to describe and model systems that are traditionally studied by social sciences. We focus our work on a specific physical phenomena: nonlinear scaling. Nonlinear scaling relations are ubiquitous in nature, appearing in a wide variety of biological, sociological, chemical and physical systems [76, 6, 69]. The statistical nature of those systems makes statistical physics a perfect tool to identify, characterize, and model them.

In the first part, we focused on the scaling relation between votes and money in a political campaign. The political elections are an area of active study on statistical physics as a model to study the collective behavior of large groups [10]. Probably the most studied feature is the vote distribution [44, 45, 40, 52]. Here, we have shown that this distribution is a direct consequence of the competitions among candidates and how much money each one invested in their campaign.

We have also shown the effect of the invested money on the final number votes. Our statistical investigation over a large number of elections shows an unambiguous correlation between these two quantities. To our surprise, the correlation revealed a sublinear scaling, which means that the average "price" of one vote grows as you increase the number of votes. In Economics this phenomena is called diseconomy of scale.

In order to understand the nonlinear scaling between the number of votes and money, we proposed a mean field model. The set of differential equations simulates a fully connected network where at every time step each candidate spends a fixed amount of money to try to convince a randomly selected voter to vote for him. We analytically solved the equations and shown that despite the simplicity, the model proved to be capable of reproducing the great complexity of the dataset. The model allowed us to fit the data using only one parameter, Δm, which is interpreted as the investment cost per vote.

The distribution of votes is well know for a large number of countries, but in this thesis we only used the Brazilian data. In the future we aim to extensively validate

our model using data from other countries. This validation process not only study the distribution of votes but also the total number of valid votes. There is some previous evidences that corroborate with our heuristic for $\approx$ %63 of turnout [57], but also other values are found [60]. It is necessary to make a detailed research in order to understand where the constant $1 - e^{-1}$ will appear and how and why deviations are found.

In the second part of the thesis we addressed a specific type of scaling, called allometry. The allometric scaling was originally proposed in biology by Kleiber [5] and extended later by others [76], a single power law comprising 27 orders of magnitude can associate the metabolisms of a microscopical bacteria and a blue whale, weighting a few picograms and more than 100 tons, respectively. Latter the same idea was applied on urban systems, but using the population of cities as analogous of mass in biology.

We studied how three behavioral urban indicators, namely the number of homicides, victims of vehicle crashes, and number of suicide scale with the population of cities in Brazil and the US. We found a superlinear scaling relation for the number of homicides with an allometric exponent that lies between 1.21 and 1.31 for different years. For the fatalities in traffic accidents we found an approximate isometric (linear) scaling. Perhaps, the most striking result is the sublinear scaling behavior between the number of suicides and city population, with allometric power-law exponents $\beta = 0.836 \pm 0.009$ and 0.870 ± 0.002, for cities in Brazil and the US, respectively. The fact that the frequency of suicides is disproportionately small for larger cities reveals a surprisingly beneficial aspect of living and interacting in larger and more complex social networks.

A complete theory of urbanization that explain allometry stills an open question [24]. One of the main problems is the lack of a fundamental definition of cities, which can sometimes lead to different exponents for different definitions [73, 27]. A complete model needs to have an unambiguous definition of cities and correct predict the power law exponent. As a perceptive we hope to adapt our election model in order to simulate the linear and sublinear allometric scaling.

REFERÊNCIAS

[1] BALL, P. The physical modelling of society: a historical perspective. *Physica A: Statistical Mechanics and its Applications*, **314.1**, 1–14, (2002).

[2] BALL, P. *Critical mass: How one thing leads to another*. Macmillan, (2004).

[3] MELO, H. P. M., MOREIRA, A. A., MAKSE, H. A., & ANDRADE, J. S. Statistical Signs of Social Influence on Suicides. *Scientific reports*, **4**, (2014).

[4] SCHMIDT–NIELSEN, K. *Scaling: why is animal size so important?*. Cambridge University Press, Cambridge, (1984).

[5] KLEIBER, M. Body size and metabolic rate. *Physiological Reviews*, **27**, 511–541, (1947).

[6] BETTENCOURT, L. M., LOBO, J., HELBING, D., KÜHNERT, C., & WEST, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. of Sci. USA*, **104**, 7301–7306 (2007).

[7] QUETELET A. Recherches sur le penchant au crime aux différens ages. *Nouveaux mémoires de l'académie royale des sciences et belles-lettres de Bruxelles*, **7**, 136 (1835).

[8] MAXWELL, J. C. *Molecules, in: W.D. Niven (Ed.), The Scientific Papers of James Clerk Maxwell*. **2**, Cambridge, p. 374, (1890).

[9] BERGER, A. L., PIETRA, V. J. D., & PIETRA, S. A. D. A maximum entropy approach to natural language processing. *Computational linguistics*, **22.1**, 39–71, (1996).

[10] CASTELLANO, C., FORTUNATO, S., & LORET, V. Statistical physics of social dynamics. *Reviews of modern physics*, **81.2**, 591, (2009).

[11] YAKOVENKO, V. M. Econophysics, statistical mechanics approach to. *Encyclopedia of Complexity and Systems Science*. Springer New York, 2800–2826, (2009).

[12] HARTE, J. *Maximum entropy and ecology: a theory of abundance, distribution, and energetics*. Oxford University Press, (2011).

[13] NEWMAN, M. E. J., & BARKEMA, G. T. Monte Carlo Methods in Statistical Physics chapter. 1–4. *Oxford University Press*. New York, USA, (1999).

[14] HUANG, K. Introduction to statistical physics. *CRC Press*, (2009).

[15] SHANNON, C. E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, **5.1**, 3–55, (2001).

[16] JAYNES, E. T. Information theory and statistical mechanics. *Physical review*, **106.4**, 620, (1957).

[17] COMTE, A. *System of positive philosophy*. Paris: Bachelier, (1830).

[18] LIND, M. *Prospect December*. p. 20, (2001)

[19] DRAGULESCU, A., & YAKOVENKO, V. M. Statistical mechanics of money. *The European Physical Journal B-Condensed Matter and Complex Systems*. **17.4**, 723–729, (2000).

[20] MANTEGNA, R. N., & STANLEY, H. E. Scaling behavior in the dynamics of an economic index. *Nature*, **376**, 46, (1995).

[21] STANLEY, H. E., AMARAL, L. A. N, CANNING, D. Econophysics: Can physicists contribute to the science of economics?. *Physica A: Statistical Mechanics and its Applications* **269.1**, 156–169, (1999).

[22] BOUCHAUD, J. P., & MÉZARD, M. Wealth condensation in a simple model of economy. *Physica A: Statistical Mechanics and its Applications*, **282.3**, 536–545, (2000).

[23] BURDA, Z., JOHNSTON, D., JURKIEWICZ, J., KAMINSKI, M., NOWAK, M. A., PAPP, G., & ZAHED, I. Wealth condensation in Pareto macroeconomies. *Physical Review E*, **65.2**, 026102, (2002).

[24] LOUF, R. Wandering in cities: a statistical physics approach to urban theory. *arXiv* preprint arXiv:1511.08236 (2015).

[25] BETTENCOURT, L. M. The origins of scaling in cities. *Science*, **340**, 1438–1441, (2013).

[26] KOZLOWSKI, J., & KONARZEWSKI M. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant?. *Functional Ecology*, **18.2**, 283–289, (2004).

[27] LOUF, R., & BARTHELEMY M. Scaling: lost in the smog. *arXiv* preprint arXiv:1410.4964, (2014).

[28] DIAMOND L., *The Spirit of Democracy: The Struggle to Build Free Societies Throughout the World* Holt Paperbacks, New York, (2007).

[29] http://elections.nytimes.com/2012/campaign-finance

[30] STRATMANN T., Some talk: money in politics: a (partial) review of the literature. *Public Choice*, **124**, 135–156, (2005).

[31] JACOBSON, G., C., The Effects of Campaign Spending in Congressional Elections. *American Political Science Review*, **72**, 769–83, (1978).

[32] ERIKSON, R. S., AND PALFREY, T. R. Equilibria in campaign spending games: Theory and data. *American Political Science Review*, 595–609, (2000).

[33] FINKEL, S. E. Reexamining the 'minimal effects' model in recent presidential campaigns. *The Journal of Politics*, **55.01**, 1–21, (1993).

[34] KRASNO, J. S., & GREEN, D. P. Preempting quality challengers in House elections. *The Journal of Politics*, **50(04)**, 920–936 (1988).

[35] WEST, D. M. *Air wars: Television advertising and social media in election campaigns, 1952-2012.* Sage, (2013).

[36] ESSER, F., & PFETSCH, B. Comparing political communication: Theories, cases, and challenges. *Cambridge University Press*, (2004).

[37] MCAFEE, R. P., & MCMILLAN, J. Organizational diseconomies of scale. *Journal of Economics & Management Strategy*, **4.3**, 399–426, (1995).

[38] RINGELMANN, M. Recherches sur les moteurs animés: Travail de l'homme, *Annales de l'Institut National Agronomique*, **12**, 1–40, (1913).

[39] http://www.tse.gov.br/

[40] CALVÃO, A. M., CROKIDAKIS N., & ANTENEODO C. Stylized facts in Brazilian vote distributions. *PloS one*, **10.9**, e0137732, (2015).

[41] JOHNSTON, R. G. C., BRADY, H. E. *Capturing Campaign Effects*. The University of Michigan Press, (2006).

[42] HOLBROOK, T. M. *Do Campaigns Matter?.* SAGE, (1996).

[43] JOHNSTON, R., PATTIE, C. How much does a vote cost? Incumbency and the impact of campaign spending at English general elections. *Journal of Elections, Public Opinion and Parties*, **18(2)**, 129–152, (2008).

[44] COSTA FILHO, R. N., ALMEIDA M. P., ANDRADE, J. S., & J. E. MOREIRA, Scaling behavior in a proportional voting process. *Phys. Rev. E*, **60**, 1067, (1999).

[45] MOREIRA A. A., PAULA D. R., COSTA FILHO, R. N., & ANDRADE, J. S. Competitive cluster growth in complex networks. *Phys. Rev. E*, **73**, 065101, (2006).

[46] HILLYGUS, D. S., & SIMON J. Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy. *American Journal of Political Science*, **47**, 583–596, (2003).

[47] GERBER, A. S. Does Campaign Spending Work?. *American Behavioral Scientist,* **47**, 541–574, (2004).

[48] CASTELLANO, C., FORTUNATO, S., VITTORIO L., Statistical physics of social dynamics. *Reviews of Modern Physics*, **81**, (2009).

[49] STANLEY, H. E. *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press, New York (1971).

[50] D. L. TURCOTTE, *Fractals and Chaos in Geology and Geophysics*. Cambridge University Press, New York, (1992).

[51] SUKI B., BARABÁSI, A. L., HANTOS Z., PETÁK F., & STANLEY, H. E. Avalanches and power-law behavior in lung inflation. *Nature*, **368**, 615 (1994).

[52] FORTUNATO, S., & CASTELLANO, C. Scaling and universality in proportional elections. *Physical Review Letters*, **99(13)**, 138701, (2007).

[53] SULLIVAN, A., & STEVEN M. S. *Economics: Principles in Action. Upper Saddle River* NJ: Pearson Prentice Hall, p. 157, (2003).

[54] BORGHESI, C., & BOUCHAUD, J. P. Spatial correlations in vote statistics: a diffusive field model for decision making. *Eur. Phys. J. B*, **75**, 395–404, (2010).

[55] KLIMEK, P., YEGOROV, Y., HANEL, R., & THURNER, S. Statistical detection of systematic election irregularities. *Proc. Natl. Acad. Sci.*, 109, 16469–16473, (2012).

[56] ARARIPE, L. E., COSTA FILHO, R. N., HERRMANN, H. J., ANDRADE, J. S. Plurality Voting: the Statistical Laws of Democracy in Brazil. *Int. J. Mod. Phys.C*, **17**, 1809–1813, (2006).

[57] TREISMAN, D. Elections in Russia, 1991-2008. *arXiv*, preprint arXiv:1605.05545 (2016).

[58] MANTOVANI M. C., RIBEIRO H. V., MORO M. V., PICOLI JR. S., MENDES R. S. Scaling laws and universality in the choice of election candidates. *Europhysics Letters* **96**, 48001, (2011).

[59] CHATTERJEE A., MITROVIĆ M., FORTUNATO S. Universality in voting behavior: an empirical analysis. *Scientific Reports 3*, **1049**, (2013).

[60] BORGHESI, C., RAYNAL, J. C. & BOUCHAUD, J. P. Election turnout statistics in many countries: Similarities, differences, and a diffusive field model for decision-making. *PLoS one*, **7**, e36289, (2012).

[61] ANDRESEN, C. A., HANSEN, H. F., HANSEN, A., VASCONCELOS, G. L., & ANDRADE, J. S. Correlations between political party size and voter memory: A statistical analysis of opinion polls. *Physica A*, **19**, 1647–1658, (2008).

[62] MEADE N., & ISLAM T. Modelling and forecasting the diffusion of innovation: A 25 year review. *International Journal of forecasting*, 519–545, (2006).

[63] MOTULSKY, H., & CHRISTOPOULOS, A. *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting.* OUP USA, (2004).

[64] ATTNEAVE, F. Applications of information theory to psychology: A summary of basic concepts, methods, and results. New York: Holt, (1959).

[65] YAKOVENKO, V. M., & ROSSER JR., J. B. Colloquium: Statistical mechanics of money, wealth, and income. *Reviews of Modern Physics*, **81.4**, 1703, (2009).

[66] MANTEGNA, R. N., & STANLEY, H. E. *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press, (1999).

[67] DURKHEIM, E. *Le suicide: étude de sociologie*, Félix Alcan, Paris, (1897).

[68] ALVES, L. G., RIBEIRO, H. V., LENZI, E. K., & MENDES, R. S. Distance to the scaling law: a useful approach for unveiling relationships between crime and urban metrics. *PLoS One*, **8**, e69580 (2013).

[69] STANLEY, H. E. *Introduction to phase transitions and critical phenomena.* Oxford University Press, New York, (1987).

[70] KLEIBER, M. *The fire of life: an introduction to animal energetics.* John Wiley, New York, (1961).

[71] BETTENCOURT, L. M., LOBO, J., STRUMSKY, D., & WEST, G. B. Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PloS One*, **5**, e13541, (2010).

[72] BETTENCOURT, L. M., & WEST, G. A unified theory of urban living. *Nature*, **467**, 912–913, (2010).

[73] OLIVEIRA, E. A., ANDRADE JR., J. S., & MAKSE, H. A. Large cities are less green. *Sci. Rep.* **4**, 4235, (2014).

[74] Brazil's Public healthcare System (SUS), Department of Data Processing (DATASUS). Available: http://www.datasus.gov.br/. Accessed 2012 Jun 1.

[75] FOWLER, J. H. & CHRISTAKIS, N. A. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *Br. Med. J.* **337**, 23–27, (2008).

[76] WEST, G. B., BROWN, J. H., & ENQUIST, B. J. A general model for the origin of allometric scaling laws in biology. *Science*, **276**, 122–126, (1997).

[77] CHANGIZI, M. A., MCDANNALD, M. A., & WIDDERS, D. Scaling of differentiation in networks: Nervous systems, organisms, ant colonies, ecosystems, businesses, universities, cities, electronic circuits, and Legos. *J. Theor. Biol.* **218**, 215–237, (2002).

[78] HILL, ALISON L., et al. Emotions as infectious diseases in a large social network: the SISa model. *Proc. R. Soc. B*, **277**, 3827–3835, (2010).

[79] NADARAYA, E. A. On Estimating Regression. *Theor. Probab. Appl.* **9**, 141–142, (1964).

[80] WATSON, G. S. Smooth regression analysis. *Sankhya Ser. A* **26**, 359–372, (1964).

[81] MONTOGOMERY, D. C., & PECK, E. A. *Introduction to Linear Regression Analysis.* Wiley, New York, (1992).