Corruption and complexity: a scientific framework for the analysis of corruption networks

Issa Luna-Pla and José R. Nicolás-Carlock[*]

* Correspondence: jnicolas@unam.mx
Institute of Legal Research, National Autonomous University of Mexico, Mexico City, Mexico

Abstract

According to United Nations, corruption is a systemic and adaptive phenomenon that requires comprehensive and multidisciplinary approaches for its effective prevention and combat. However, traditional approaches lack the analytical tools to handle the structural and dynamical aspects that characterize modern social, political and technological systems where corruption takes place. On this matter, complex systems science has emerged as a comprehensive framework to study highly adaptive phenomena from natural to socio-technical settings. Thus, in this article we present an empirical approach to model corruption using the concepts and tools of complexity science, mainly, complex networks science. Under this framework, we describe a major corruption scandal that took place in Mexico involving a network of hundreds of shell companies used to embezzle billions of dollars. We describe the structure and dynamics of this corporate network using available information related to their personnel and the date of the companies' creation. We measured some global parameters, such as density, diameter, average path length, and average degree in order to provide systematic evidence on which corporate characteristics are likely to signal corruption. Moreover, this analysis also provides an objective perspective of the systemic nature of events where companies are abused for corrupt purposes, and the shortcomings of reductionistic analyses. Major corruption scandals comprise both legal and illegal deeds, in addition to several parties acting simultaneously over extended time periods. As a whole, such scandals pose enormous challenges for the study of law and put the legal design of administrative and criminal controls to the test.

Keywords: Complex networks, Complex systems, Complexity, Corruption, Corruption networks

Introduction

> *"Injustice anywhere is a threat to justice everywhere. We are caught in an inescapable network of mutuality, tied in a single garment of destiny. Whatever affects one directly, affects all indirectly."*
> *- Martin Luther King Jr.*

"The 21st century will be the century of complexity."
- Stephen Hawking

The purpose of modern governments is to establish and enforce the rules that guarantee social cohesion, personal freedom, and collective well-being. In contrast, corruption comprises everything that deviates from that purpose by distorting the goal for which all socio-political structures are created, sacrificing the well-being of the collective for the benefit of the few.

According to the UN Convention Against Corruption (UN General Assembly, 2003), corruption is no longer acknowledged as a local problem of the nations, but as a transnational phenomenon that affects societies in deep and several ways (Transparency International, 2018): on the political front, corruption is an obstacle to the development of democracies and the rule of law, affecting the political leadership and institutional legitimacy; in economy, corruption hinders growth and distorts healthy competition within the markets, deterring national and foreign investments; at the ecological level, corruption degrades the environment, destroying vital ecosystems through the reckless, unchecked exploitation of natural resources, with clear local and global consequences; moreover, corruption corrodes the fabric of society by generating environments that foster the violation of human rights. In summary, corruption produces a complex, ubiquitous and many-faceted threat to the common interests of all societies that allows for enormous systemic risks in different sectors, form local to global levels (Helbing, 2013).

It is well acknowledged that the theoretical and technical frameworks of traditional corruption studies are not sufficient in order to handle the highly systemic nature of this phenomenon, and that new empirical, inter-disciplinary, and scientific approaches are necessary if we are to face the complexity of corruption effectively (Mungiu-Pippidi, 2017). At this juncture, issues relating to legal matters find common ground with problems of a scientific nature considering that the purpose of science, technology and innovation is to study our natural, social and technological environments, aiming at a deeper understanding of such environments and the application of the obtained knowledge in order to improve our living standards (Pentland, 2015; Altshuler & Pentland, 2018). In particular, through the transformation and advancement of legal practices relevant to that goal (Livemore & Rockmore, 2019). In this regard, complexity science has emerged as a comprehensive framework that allows for the multidisciplinary study of natural and social adaptive systems that are found everywhere in our everyday lives (Ball, 2003; Mitchell, 2009; Vespignani, 2012; Thurner, Hanel & Klimek, 2018).

In this article we present an empirical approach to describe and model corruption using the concepts and tools of complex systems science, mainly, network science (Barabási, 2016; Newman, 2018; Thurner et al., 2018). Under this framework, we describe a major corruption scandal that took place in Veracruz, Mexico, between 2010 and 2016, involving a complex network of hundreds of shell companies used to embezzle billions of dollars, originally destined to diverse social programs (Animal Político, 2016). First, we survey the progress of traditional studies on this topic by briefly reviewing attempts at describing and fighting corruption from different academic perspectives. Second, we present a short introduction to the characteristics of complex systems. Finally, we apply the complexity and network approaches to the analysis of the Veracruz case. Here focus on the structural and dynamical features of this corporate network in order to identify

the network characteristics that are more likely to signal corruption. Also, we show that corruption networks go beyond the rhetorical elements of public discourse and practice, materializing into well-defined and complex structures across different layers of information (for example, through legal representatives or administrators) and time periods (date of creation of the companies), providing an objective view of the systemic nature of this event, and the shortcomings of reductionistic investigations.

Define, measure, predict and control

From a scientific perspective (Barabási, 2010; Vespignani, 2012), modern studies on corruption should aim at defining, measuring, and predicting the phenomenon, in a way that the mechanisms and methods for its control (regulation) and elimination (combat) may be set and executed satisfactorily. In science, these objectives set the conceptual and methodological frameworks that allow results to be efficiently and effectively implemented, otherwise, the effectiveness of the methods is wanting and in a best-case scenario the results are temporarily adequate, and in the worst-case, they may be counterproductive (Milinski, 2017; Muthukrishna, Francois, Pourahmadi & Henrich, 2017). Considering this, the complexity of anti-corruption efforts lies first and foremost in the adequate development of the following: (i) definition, (ii) measurement, (iii) prediction, and (iv) control. These goals not only organize the strategies that must be implemented for the study of the phenomenon, but also, they provide a guideline for the appraisal of previous anti-corruption approaches, their limitations and their potential for future improvement.

Defining corruption

A quick check to the literature allows us to see, there is a large amount of studies addressing the conceptualization or definition of corruption (Andvig et al., 2001; Riccardi & Sarno, 2014). Several definitions for corruption have been proposed in proportion to the number of social, economic and political areas were corruption has taken place (definitions according to sector), as well as from the prism of the multi-discipline (definitions according to discipline). For example, a large portion of the literature discusses the definition of corruption from economic, legal and governmental perspectives, aiming to create a definition drawn from the "umbrella idea", that contains concepts such as clientelism, abuse of power, state capture, and patrimonialism (Varraich, 2014). One may also find definitions of corruption in government and bureaucracy (Rose-Ackerman & Palifka, 1999), political corruption (Heidenheimer & Johnston, 2011; Reno, 1995), and the definition of corruption as moral decadence and lack of ethics (Huntington, 1970; Mulgan, 2012). Similarly, social research has focused on classifying the types and sub-types of corruption in order to be operational. For example, several categories classify corruption based on its economic magnitude with concepts such as grand corruption, structural corruption, systemic or endemic corruption, petty corruption (Andvig et al., 2001), and rentism (Khan & Jomo, 2000). It is noteworthy that the most popular definition is the one promoted by Transparency International and the World Bank that states that corruption is "the abuse of entrusted power for private benefits or gains" (Transparency International, 2018). This definition is influenced by non-moralistic classic concepts (Nye, 1967), as well as by the historical endeavor of philosophy to search for the causes and origins of corruption (Hill, 2012).

Even though the efforts to define corruption from general to specific settings are diverse, one may argue that those definitions have little to do in current anti-corruption efforts since at the end of the day, the legal definitions of corruption, such as bribery, embezzlement or obstruction of justice (UN General Assembly, 2003) are the ones against whom prosecutors and investigators can do something about it, in other words, where there's no crime there's no punishment. Nonetheless, the task of defining corruption is important because it bounds the subject matter, however, this task must go beyond the mere delimitation of the phenomenon in different contexts, but also strive to generate an operational and quantifiable definition, so that its causes and possible effects may the described with greater precision (Olken & Pande, 2012; Riccardi & Sarno, 2014).

Measuring corruption

Over the past 20 years, empirical sources of information have been used to have a better and objective approximation to the reality of the corruption phenomenon by establishing proxies, risk indicators, correlations among relevant indexes and socio-economic parameters (Lambsdorff, 2007; Olken & Pande, 2012). Within this academic current one may find several indexes that try to measure corruption in a standardized manner by quantifying risk levels based on experience and perceptions (Svensson, 2005). One may also find the economic behavior and anthropological research that studies honesty, cooperation, and reciprocity (Drugov, Hamman & Serra, 2014; Sah & Loewenstein, 2014; Arney, 2010). In addition, economic theories strive to explain the way in which corruption occurs by using game theory, conceptual models of transactions or trade, and moral dilemmas (Platteau, 1994; Yoo, 2008).

Even though these studies represent progress in the modern description and quantification of corruption, they still suffer from important shortcomings. For example, the interpretation of perception indexes is highly dependent on the professional expertise of the group sample (Morris, 2018), and suffer from conceptual limitations that tend to over-simplify the phenomenon while showing inconsistencies among results when compared to each other over different periods of time (Méndez & Sepúlveda, 2009). It is also well-known that these approaches are based on statistical models and socio-economic parameters that do not allow to establish causality in a precise manner, nor are they able to describe corruption acts from their micro- to their macro-properties (Mungiu-Pippidi, 2017). Nevertheless, important efforts have been conducted in order to establish objective risk indicators, for example, in public procurement (Fazekas, Tóth & King, 2016; Fazekas & Kocsis, 2017).

Predicting corruption

The possibility of establishing models that can be used to quantify corruption provides an opportunity not only to describe the phenomenon more precisely, but also to predict it. However, the problem of prediction is not rigorously addressed within the vast academic literature about corruption, and we know little of the aspects that may help predict these acts in a precise manner (Colonnelli, Gallego & Prem, 2019). For example, the statistical models used in diverse studies (Olken & Pande, 2012) lack the capacity of prediction due to the scope of the data used, the spatial and temporal scales of analysis, the fields of operation, and the high dependency on correlations, making nearly

impossible to establish causalities for different events (Riccardi & Sarno, 2014). As it is, conclusions and proposals derived from such models should be considered reservedly within their respective contexts. Nevertheless, with current computing capabilities and trends towards the digitalization of information in many public sectors, the ideal and promising approaches to predict corruption risk seem to be those supported by live-monitoring systems powered by artificial intelligence (Fazekas & Kocsis, 2017; López-Iturriaga & Sanz, 2018; Colonnelli et al., 2019). Such systems may provide a wide range of real-time preventive capabilities (Byers, 2017; Altshuler & Pentland, 2018).

Controlling corruption

The study of corruption has helped to delve deeper into practical cases and to have a better understanding of the workings of sophisticated operational schemes by identifying the common traits of corrupt acts, types of corrupt human behavior, normative spaces and grey areas in law that might allow for corruption (Andvig et al., 2001). Solutions for fighting and controlling corruption proposed by these studies consist of legal reforms, novel institutional and burocratic designs, mechanisms for accountability and control, codes of ethics, awareness campaigns, among other strategies for legal combat, prevention and sanction (Riccardi & Sarno, 2014; David-Barrett, Fazekas, Hellmann, Mark & McCorley, 2018). Great efforts have been made to improve the measurement of corruption, however, these suffer from limitations in their design (Méndez & Sepúlveda, 2009); studies recommending solutions for corruption that consist of more rules and legal reforms are under debate due to the fact that ineffectiveness and negative impacts from overregulation may arise (Mungiu-Pippidi, 2017; Mungiu-Pippidi & Dadašov, 2017; Smilov, 2010). With current scandals of corruption worldwide, the results obtained thus far remain highly questionable and the criticism concerning the scientific progress in corruption studies is ever more present.

In summary, appropriate definitions of corruption would make objective measurements not only possible, but also more accurate. In turn, the correct measurement would make way for the possibility of prediction, and prediction would give rise to the intelligence needed for control. Though great progress has been made in the fight against corruption, all of these efforts suffer from the inherent limitations pertaining to the methodologies that support them. Having exact models of the causes, circumstances, mechanisms, effects and consequences of a corruption event remains as a great scientific challenge. Without a doubt, we know much more about corruption than we did 20 years ago, yet, this issue remains unsolved: we do not know how to measure it in a systematic and standardized manner, and we still have not been able to control it within existent models. It is clear that this problem demands of new conceptual, methodological and analytical frameworks that allow for the integration of all the elements that in one way or another determine the complexity of the phenomenon.

Applied complexity science

Corruption is a phenomenon that occurs within the intricate structure of social, political and technological systems, therefore, a better understanding of it requires of integral and inter-disciplinary perspectives. As with contemporary medicine or biology advances, which stand upon the progress made by disciplines such as physics, applied mathematics

and computing science, social sciences have begun to use such disciplines to describe, model, explain, and even predict certain phenomena (Conte et al., 2012; Holme & Liljeros, 2015; Lagi, Bar-Yam, Bertrand & Bar-Yam, 2015; Wiesner et al., 2018; Capraro & Perc, 2018). This integration of disciplines previously considered totally apart from each other (natural and social sciences, and even the humanities) has been possible due to important advances in computing capabilities, knowledge-transfer and cross-disciplinary problem solving (Ball, 2003; Miller & Page, 2009; Caldarelli, Wolf & Moreno, 2018).

Moreover, in the scientific exploration of physical to social systems, it has been found that the hardest adaptive systems to model and control are those involving individuals whose willing decisions may give rise to collective phenomena that are not easily defined, explained or predicted by means of the analysis of isolated individuals (Ball, 2003; Miller & Page, 2009; Caldarelli et al., 2018; Capraro & Perc, 2018). In this context, corruption could be regarded as a phenomenon that occurs within systems whose structure and dynamics can evolve as a response to changes in its corresponding socio-political and regulation context, with a strong dependence on the interrelation of different factors and actors acting as a whole. In general, systems with the previous characteristics are the subject matter of complexity or complex systems science, which represents a new scientific paradigm and a new way of doing science in the twenty-first century (Mitchell, 2009; Thurner et al., 2018; De Domenico et al., 2019; Helbing et al., 2015).

Complex phenomena

In complex systems theory, a system is considered complex not only because it has an intricate structure, but also because its temporal evolution cannot be easily explained as a function of the behavior of its isolated components (Bar-Yam, 1997; De Domenico et al., 2019). In particular, there are two important concepts that help to explain the 'complexity' of a system, those are 'emergence' and 'self-organization' (Sayama, 2015). On the one hand, 'emergence' is a concept associated to the effects of non-trivial interrelations among the components of a system across different scales of observation or analysis. Specifically, the properties of the parts acting as a whole are called 'emergent' when they cannot be explained based on the properties of the parts looked in isolation, thus, global properties are different from local ones (Bar-Yam, 1997; Sayama, 2015). On the other hand, 'self-organization' is a dynamic or temporal process through which the solely interactions among the multiple parts of the system create collective structures and behaviors, with no intervention from a central or external organizing agent (Sayama, 2015). From these concepts, it becomes clear that there are two important aspects to be considered within the analysis of complex systems: the structure (statics) and temporal evolution (dynamics) of the system. In complexity science, network theory is one of the most important tools for the analysis of these structural and dynamical elements (Sayama, 2015; Thurner et al., 2018).

Complex networks

Network theory has been applied to have a better understanding of many natural, socio-technical, and legal systems (Barabási, 2016; Rutherford, Lupu, Cebrian, Rahwan, LeVeck & Garcia-Herranz, 2018). The importance of the application of multidisciplinary and scientific approaches such as complex systems, network theory, and even physics, to the

study of criminal activities was presented by the end of last century (Sparrow, 1991). However, it was until the last decades that these types of studies have begun to gain momentum given the great progress in computing and data science (Caldarelli et al., 2018) and their enormous relevance in modern social, economic, and political contexts (D'Orsogna & Perc, 2015; Helbing et al., 2015; Espinal-Enríquez & Larralde, 2015; Marshak, Rombach, Bertozzi & D'Orsogna, 2016; DellaPosta, 2017; Fazekas, Skuhrovec & Wachs, 2017; Morselli & Boivin, 2017; Altshuler & Pentland, 2018; Magliocca et al., 2019; Ouellet, Bouchard & Charette, 2019; Niu, Elsisy, Derzsy, & Szymanski, 2019).

Notably, although corruption studies go back a long way and have been conducted from different perspectives, corruption studies conducted from a complex systems or network theory approaches are quite recent and scarce. For instance, a recent study covering 30 years of corruption in Brazil shows that the co-occurrence of politicians in corruption scandals create networks with large connected components that might spans decades (Ribeiro, Alves, Martins, Lenzi & Perc, 2018). Other study proposes diverse methods for the strategic dismantling of corruption or crime networks considering their structure and the cost of removing key nodes (Ren, Gleinig, Helbing & Antulov-Fantulin, 2019). Another study looks into the social fabric of Hungary in order to establish the social factors associated to corruption risk in public procurement, finding that fragmented social networks are more prone to corruption risk whilst more diversity hinders it (Wachs, Yasseri, Lengyel & Kertész, 2019). An additional study that explores 28 years of bill-voting in Brazil shows that the dynamics of co-occurrence networks of similar-voting congressmen reveal patterns that allow for the identification of convicted corrupt politicians and also, for the possibility of predicting or identifying other possible corrupt individuals within the network (Colliri & Zhao, 2019). Noteworthy, the identification of latent criminal groups (Campedelli, Cruickshank & Carley, 2019) and the effective dismantling of their organizational structure (Wandelt, Sun, Feng, Zanin & Havlin, 2018) are relevant and non-trivial subjects in criminal investigations and law enforcement, since empirical evidence has shown that the dismantling process might potentially make these criminal organizations stronger (Duijn, Kashirin & Sloot, 2014). In addition, when it comes to fighting corruption the goal is clear: one not only is looking to describe it post factum, but to predict it (Rumi, Deng & Salim, 2018; Alves, Ribeiro & Rodrigues, 2018; López-Iturriaga & Sanz, 2018; Colonnelli et al., 2019; Wachs et al., 2019; Wachs & Kertész, 2019; Colliri & Zhao, 2019).

In summary, though corruption studies are diverse and tackle different aspects of the phenomenon, complex systems and network science approaches allow us to establish practical aspects for its investigation (Sayama, 2015; De Domenico et al., 2019), mainly:

I. *Components and interactions. Complex systems are usually comprised of large sets of interacting elements or components. Both the components and their interactions may be of different types.*

II. *Network structure. The structure of a complex system may be described as a network of interactions and interrelations in which all nodes and edges evolve as a whole over time.*

III. *Self-organization. There are many interactions among components. These occur independently, with no need of intervention by central organizing agents, producing collective non-trivial phenomena.*

IV. *Emergence. The collective behavior and properties of these systems can neither be predicted nor understood from the individual behavior or properties of each component in isolation.*

V. *Predictability and control. The dynamics, or temporal evolution, of a complex system are collective and often non-linear which, under certain conditions, makes the system highly unpredictable and difficult to control.*

To show the usefulness of approaching the study of corruption through the complexity perspective, we will now analyze one of the most important corruption scandals in Mexico of the past decade.

Corruption networks

In the Mexican case, recent corruption scandals tell of the complexity of the phenomenon. From governors accused of embezzlement at local levels (Ángel, 2017), through federal agencies and public universities awarding contracts to shell companies at a national scale (Roldán et al., 2018), to transnational bribery scandals (Olmos, 2018), corruption is ever present across all public and private sectors of Mexico. As a concrete example, here we analyze a recent and paradigmatic corruption scandal: the case of the "phantom" companies of former governor of Veracruz, Javier Duarte de Ochoa (Ángel, 2017; Animal Político, 2016).
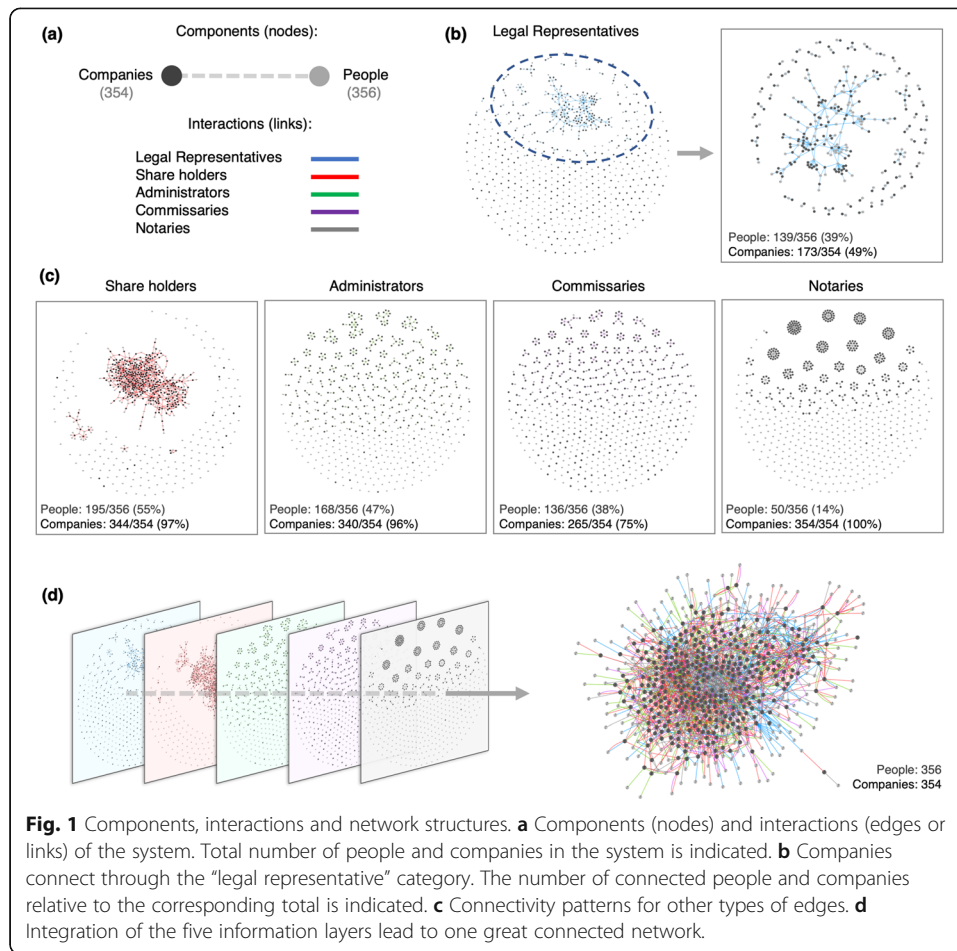
Case, data and methodology

In one of the biggest embezzlement scandals in Mexico of recent times, the former governor of the state of Veracruz, Javier Duarte de Ochoa was sentenced by federal authorities to 9 years of prison for criminal conspiracy and money laundering, after an initial charge for organized crime and operations with illicit proceeds, and an unexpected escape and extradition from Guatemala (Animal Político, 2016). According to the Mexican Superior Audit of the Federation, the amount of diverted money could reach 60 billion pesos (about 3 billion dollars), originally destined to diverse social programs, security and education, but that never reached their end. As the leading investigative journalists of the case documented (Ángel, 2017; Animal Político, 2016), the embezzlement mechanism consisted of hundreds of "phantom" or shell companies, created under fake ownership, that were awarded with contracts from the local government for diverse projects but that never delivered neither the goods nor services promised. This

is a case in which much of the analysis and debate centers on the legal fulfillment of the public procurement process, the failing of the control mechanisms, and the government responsibility on these matters. However, the analysis of the abuse of companies for corrupt practices is a relevant aspect at the forefront of international anti-corruption efforts (Fazekas & Tóth, 2017). This matter has not been formally treated for the Duarte case and it is especially important for two reasons: first, the lack of systematic evidence on which corporate characteristics are likely to signal corruption has the potential to bias our understanding of corruption, making it overly focused on the public sector (Fazekas & Tóth, 2017); second, the Duarte case is eminently characterized by a huge group of private companies that seem to be structured to operate as a network. Recent studies have shown that the characterization of the structural and operational features of corporate or trade networks are not simple, with different layers of information adding to the complexity of the problem (Alves et al., 2019; de Jeude, Aste & Caldarelli, 2019). Therefore, in the present analysis we focus on the description of the structural and dynamical features of the network of companies that were used in this scandal in order to find company corruption risk indicators, in particular, in the ownership and management structures.

The data used for the analysis comes from a dataset gathered from official sources (open to public access under Mexico's General Law of Transparency and Access to Information) by the NGO known as Mexicanos Contra la Corrupción e Impunidad (Mexicans Against Corruption and Impunity) and the leading group of investigative journalists of the case, known as Animal Político (Animal Político, 2016). This dataset contains information about 354 companies and 356 people associated to those companies. For legal reasons, we have held all information regarding names and official ID's of companies and people in anonymity. For each company, the dataset contains available information about their legal representatives, shareholders, administrators and commissars, as well as the notaries that formalized them. For the network analysis we considered a bipartite approach in which companies are related to people through five different types of edges: share-holders, legal representatives, administrators, commissars and notaries (see Fig. 1a). As we show below, this classification is due to the fact that there are individuals with multiple work relationships within the same company or among different companies, therefore, instead of encoding this information into the nodes, we decided to encode it into the edges and to treat each category as layer of information for this network. Thus, each layer contains the same number of companies (354) and people (356).

As previously stated, the networks do not show the government agencies that granted contracts to private companies, since the goal of this study is to explore the companies' personnel role in the structure and evolution of the network. As network metrics we considered the density (fraction of the number of real edges of a network to the number of its possible edges given the number of total nodes), diameter (maximum distance between a pair of nodes in the network), average path length (the average distance between pairs of nodes in the network), average degree (average of the number of neighbors for each node), clustering coefficient (fraction of closed triplets to the total number of closed and open triples, where a triplet is a subset of three neighboring nodes), number of connected components (number of connected sub-sets or sub-network), and number of multi-edge node pairs (in a multigraph or multi-edge network, these are the number of multi-edges

Fig. 1 Components, interactions and network structures. **a** Components (nodes) and interactions (edges or links) of the system. Total number of people and companies in the system is indicated. **b** Companies connect through the "legal representative" category. The number of connected people and companies relative to the corresponding total is indicated. **c** Connectivity patterns for other types of edges. **d** Integration of the five information layers lead to one great connected network.

connecting any pair of nodes) (Newman, 2018; Barabási, 2016). The network was analyzed using Cytoscape, a well-known open-source software for network analysis (Shannon et al., 2003).

In the following sections, we will explore the qualitative and quantitative features of Duarte's phantom-companies network through the lens of complex systems and network science: first, the components, interactions and network structures; second, the self-organizing and emergent elements; and finally, some aspects on predictability and control.

Components, interactions and network structure

The components are defined as the participating companies and people. These define two types of nodes that are linked through five categories: shareholders, legal representatives, administrators, commissars, and notaries (see Fig. 1a). As stated before, the information in such categories is encoded within the edges and not in the nodes, due to the fact that numerous people play several roles within the same company or in different companies. Under that classification, each type of edge reveals part of a complex network structure that is comprised of five layers of information and each layer contains the same number of companies (354) and people (356). For example, the "legal

representative" category produces a bipartite network that consists of a sub-set of 139 people related to 173 companies (See Fig. 1b). In this way, each category reveals different networks within the same system (see Fig. 1c). These features also reflect quantitatively in the network metrics (see Table 1).

In general, these networks have low density (not all companies connect to all people), which is characteristic of real-world networks (Barabási, 2016); they also have low average degree, which means that there is a great majority of companies/people that connect to at most another one people/company in each layer; clustering coefficient equal zero, due to the fact that these networks are bipartite by construction and therefore, closed triplets are impossible to be generated. This leaves the diameter (largest distant between pairs of nodes) and average path length (average distance between pairs of nodes), that are measures of the extent to which connected companies and people are present in the network due to the fact that there are multiple people connecting multiple companies at once. As it can be observed, the "shareholder" network is the one that generates more structure among the rest of the categories, followed by the "legal representatives", "administrators", "commissars" and finally, "notaries" (see Fig. 1 and Table 1).

From the previous results, one can easily see that "shareholders" are the ones that provide great cohesion to the network, since they provide more connectivity among the total number of companies. In context, this finding provides an important insight in the way corrupt companies might be investigated, specifically, given that not all actors have the same role within this type of systems, prosecutors must be aware that looking at just one individual or group might lead to missing details of bigger schemes. Moreover, since each type of edge constitutes a single layer of information within the system, then, their integration reveals a more exact version of the network. Remarkably, after aggregating all layers into one, we find that all nodes are condensed into one fully connected and complex multigraph or multi-edge network (see Fig. 1d and Table 1). This

Table 1 Network metrics for the integrated network decomposed by layer of information

| Parameter/Layer	Shareholders	Legal Reps.	Administrators	Commissars	Notaries	Average (SD)	Aggregated
People*	195 [55%]	139 [39%]	168 [47%]	136 [38%]	50 [14%]	137 (54)	356 [100%]
Companies*	344 [97%]	173 [49%]	340 [96%]	265 [75%]	354 [100%]	295 (77)	354 [100%]
Total nodes	710	710	710	710	710	710	710
Total edges	763	304	370	277	354	414 (199)	2068
Density	0.0030	0.0010	0.0010	0.0010	0.0010	0.0014 (0.0009)	0.006
Diameter	16	16	6	6	2	9 (6)	10
Av. path length	6.5	6.0	2.6	1.8	1.8	3.7 (2.3)	4.7
Av. degree	2.1	0.8	1.0	0.7	0.9	1.1 (0.6)	4.3
Clustering coefficient	0	0	0	0	0	0	0
Connected comp.	177	451	340	433	356	351 (109)	1

The indicated number corresponds only to connected companies and people for each layer. The percentage in brackets relative to 356 people and 354 companies in total, respectively

finding is explained by the fact that there are several people who perform multiple roles within a single company or among different companies (as explained further below).

The previous analysis through decomposition has advantages and disadvantages: on one hand, it allows to observe the details concerning the contribution of each information layer to the system's structure: in corporate networks, not all players provide the same amount of information; on the other hand, when considering a single category for analysis, highly valuable information concerning connectivity among companies is excluded. This is a very important insight that complexity approach provides: if we truly wish to understand complex systems such as the ecosystem of corporate networks in contracting markets, we must be aware that excluding information -- willing or unwillingly -- can hamper our perspective of the exact workings of the system. In our case, companies that might look separated or independent from one another are then shown to be connected, forming one fully connected network, when the information about their personnel is integrated.

Self-organization and emergence

The aggregated multigraph can be decomposed as a function, for instance, of the year of the creation of the companies (Fig. 2a, b and c). This temporal decomposition by year reveals that the large multigraph emerges from a self-organization process of subsets of companies created in different stages in time, that in turn create fully connected networks (Fig. 2b). Quantitatively, these features reflect through their metrics per year. Without loss of generality, the contributions of the initial and final year to the total
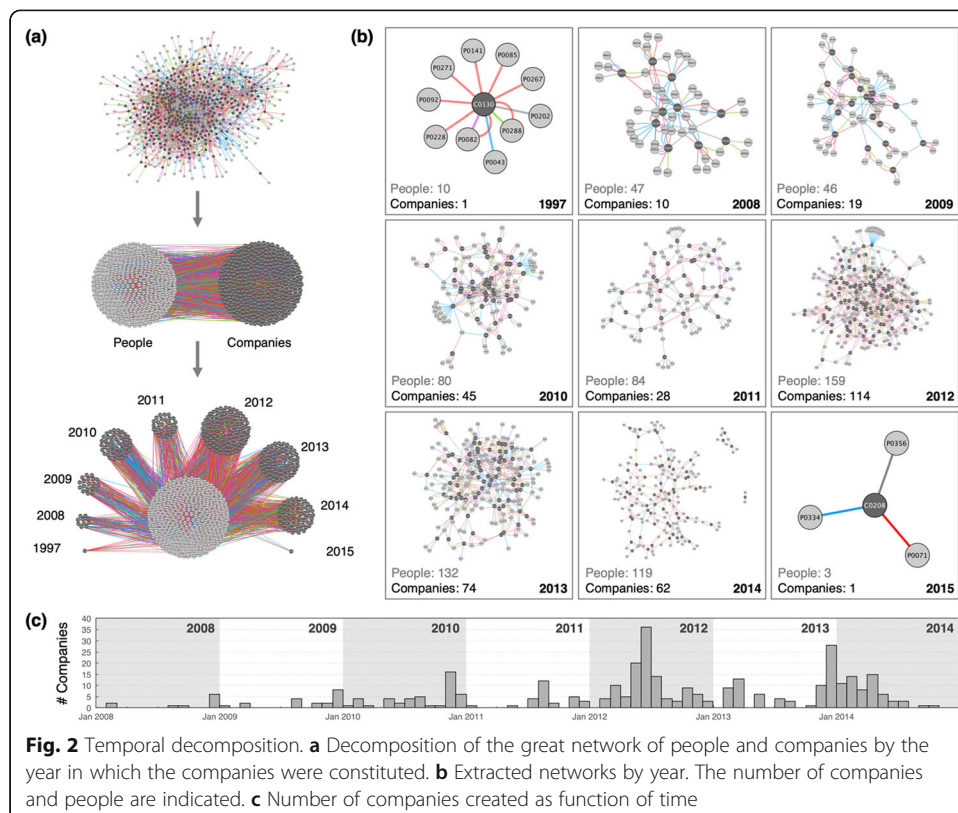
Fig. 2 Temporal decomposition. **a** Decomposition of the great network of people and companies by the year in which the companies were constituted. **b** Extracted networks by year. The number of companies and people are indicated. **c** Number of companies created as function of time

network are negligible and thus are discarded (see Table 2). In all these networks (2008–2014), both the number of nodes and edges changes over time; the number of connected components clearly shows the fact that one network of companies was created per year; the density remains low, and again, the clustering coefficient is equal to zero due to the bipartite nature of the network.

Notably, metrics such as the diameter, average path length and average degree remain almost of the same magnitude per year. In short, these metrics do not provide enough information to establish corporate corruption risk indicators. However, one can take advantage of the fact that the cohesion of the multigraph arises from the total contribution of "shareholders", "legal representatives", "administrators", "commissars", and "notaries". Here, one can measure the number of multi-edge node pairs, these are the number multi-edges connecting any pair of nodes in the bipartite multigraph. For company-nodes, multi-edge pairs correspond to the number of people with at least one role or work relationship within the company. For people-nodes, multi-edge pairs correspond to the number of companies in which a given individual has multiple roles. Remarkably, multi-edge node pairs have strong variations during the years of high creation of companies, pointing towards relevant anomalies, showing their potential as a corporate corruption risk indicator (see Table 2).

The behavior of separate companies is not the same as the behavior of companies acting as a whole. The self-organized dynamics of the corporate network of this case was only revealed through the full integration of all the information available on companies' personnel. This emergent property, i.e. the way it operates as a whole complex network, is even clearer when one considers how metrics differ from each other when performed over one layer or 1 year than when performed over the aggregated multigraph. In the structural analysis per layer or in the decomposition per year, each category has its own characteristic metrics. It would be tempting to make an average of the information to infer the behavior of the whole network, however, as we can observe from the metrics of the aggregated network, these differ from simple averages (see Table 1 and Table 2). This is due to the fact that, in complex systems, interrelations among components are non-linear and therefore, and system properties behave in the same manner as well.

Table 2 Network metrics for the integrated network decomposed by year of companies' creation

Parameter/Year	2008	2009	2010	2011	2012	2013	2014	Average (SD)	Aggregated
People[a]	47	46	80	84	159	132	119	95 (43)	356
Companies[a]	10	19	45	28	114	74	62	50 (36)	354
Total nodes	57	65	125	112	273	206	181	146 (79)	710
Total edges	94	128	273	164	628	454	312	293 (192)	2068
Total multi-edge pairs	15	36	76	27	192	114	64	75 (61)	526
Density	0.049	0.044	0.025	0.022	0.012	0.016	0.015	0.026 (0.015)	0.006
Diameter	6	8	8	10	12	10	14	10 (3)	10
Av. path length	3.6	4.1	4.4	5.4	5.2	4.8	6.4	5 (1)	4.7
Av. degree	2.8	2.8	3.1	2.4	3.1	3.2	2.7	2.9 (0.3)	4.3
Clustering coefficient.	0	0	0	0	0	0	0	0	0
Connected comp.	1	1	1	1	1	1	3	1.3 (0.8)	1

[a]The indicated number corresponds only to connected companies and people for each network

The previous results, the creation of one connected network per year and the multiplicity of edges between pairs show us that, for this network, many companies have shared the same personnel over the years, many of them with multiple roles within the same and/or among other companies, thus, reflecting a remarkable anomalous behavior in the structural and dynamical growth of this network. Also, emergent properties measured through network metrics are highly dependent on the integration of all available information across different scales of observation. Again, in complex systems, we must be aware that the temporal extension of an event like this one could reveal only a small part of the greater scheme. Excluding information -- willing or unwillingly -- can hamper our perspective of the exact workings in the system, which in our case, reveals how one fully connected network of companies is created through time.

Predictability and control

From the previous structural and dynamical analysis, we conclude that the information about the personnel associated to each company plays a key role in the description of a given event, in the definition of good corruption risk indicators, and in the possible detection of irregularities within organized schemes of network operations. Heuristically, network visualizations are a quite powerful tool that provides an intuitive base for identifying and understanding certain behaviors and patterns (Venturini, Jacomy & Jensen, 2019). However, the true power of network science does not entirely lie in its capacity to describe any particular system in a qualitative manner, but rather in the mathematical formalism that this framework gives to the definition, quantification and modeling of the different parts of the system. As we showed, network metrics can be defined either for each node or for the whole network, however, not all might be useful in order to characterize and detect relevant anomalies that point towards corrupt behaviors in corporate networks, specifically, multi-edge bipartite networks. In the analysis per layer or edge type, only the diameter and average length path reveal information that could be relevant for our purposes. In the analysis per year, it was the multi-edge node pairs, a measure of the degeneracy of a given node, the one that revealed interesting anomalies related to the multiplicity of work relations. Certainly, these insights could be used to define risk indicators and control mechanisms for companies participating in public procurement; but of course, they still need to be tested in further cases and in other datasets of corporate ecosystems, something that goes beyond the scope of the present article. Nevertheless, the results presented here show how complexity science along with network theory allows us to define, quantify and model a given corruption event, in order to identify and establish possible prediction and control mechanisms (Vespignani, 2012; Barabási, 2016; Ren et al., 2019).

Conclusions & remarks

In this article, we presented empirical evidence of a grand corruption scandal in Mexico that provides a glimpse into the complexity of events where companies are abused for corrupt practices. The analysis performed looks to fill a gap in the lack of systematic evidence on which corporate characteristics are likely to signal corruption in public procurement. To this end, we focused on the description of the structural and dynamical features of the network of shell companies of the Duarte case in order to find

company corruption risk indicators, in particular, in the ownership and management structures. This was done under a complex systems and network theory approach where companies and their personnel were model as a multi-edge bipartite network:

I.   For the structural analysis of the network (Fig. 1), we found that all the companies and their personnel created a one big connected component when the five layers of information were integrated. Among these, the "shareholder" layer was the one that provided more information about the connectivity of the network, followed by the "legal representatives", "administrators", "commissars" and finally, "notaries". For this part, the network metrics used were non-conclusive in order to establish clear corporate corruption risk indicators (Table 1).

II. For the dynamical part (Fig. 2), we found that when the bipartite multigraph was decomposed into the year of companies' creation, one connected component appeared for most years. Again, most network metrics were non-conclusive in order to establish clear corporate corruption risk indicators, except for the number of multi-edge node pairs, that provided a good measure of the degeneracy of both companies and people (Table 2). For company-nodes, multi-edge pairs correspond to the number of people with at least one role or work relationship within the company. For people-nodes, multi-edge pairs correspond to the number of companies in which a given individual has multiple roles.

III. The evidence presented here points towards the hidden complexity in other corruption events of this kind, where not all information of the parties involved is available, and the shortcomings of analysis through reductionistic eyes.

IV. Corruption is a phenomenon that not only occurs within the complexity of private or public socio-political and technical systems, but also, it can give rise to a complex system in itself, in which the interrelations of different actors and factors acting as a whole originate characteristic self-organizing and emergent phenomena.

Empirical and multidisciplinary corruption studies are key in order to design effective anticorruption strategies. From a legal perspective, the phenomenon of corruption poses great challenges, such as the detection of concurrent practices, the correction of normative frameworks, the control and prosecution of corruption crimes, the decreasing of the incentives that feed them, and the recovery of assets. The empirical study of great corruption scandals, such as the one presented here, has revealed that corrupt activities might consist of both legal and illegal deeds, from public to private sectors, and with multi-role actors that act simultaneously over long periods of time. As a whole, they go far beyond the institutional capability and legal design of administrative controls, as well as surpassing causes of

jurisdictional sanction. Nevertheless, the complexity approach has the potential to change the way in which we study and implement law. Certainly, it proves the need for improving criminal and administrative forensic investigations through the scientific innovation of studies that pertain to law and justice. In addition, great importance should be given to efforts that allow for the coordination among key actors and prosecutors, and to the exchange of information through which public and private sectors can generate relevant intelligence for the creation of effective anti-corruption networks.

Finally, complex systems science is a new paradigm for doing science as it provides a conceptual and analytical framework from which one might understand a great number of phenomena from natural to social systems. Political, economic, ecological and social systems are fine examples of complex systems that greatly impact our lives and that can be studied under the adaptive systems approach (Helbing et al., 2015). The major problems facing modern societies, such as climate change, migration, inequality, crime or corruption, may be more likely to be solved when public policies are based not only on science, but on a new kind of science, one that considers the inherent adaptive and systemic aspects of modern States, that is, applied complexity science (Geyer & Cairney, 2015). In addition, it is highly important to emphasize that the control mechanisms inferred from these types of studies will only be useful when implemented within the corresponding administrative and legal frameworks that provide them of support and operative reach (UN General Assembly, 2003). In this aspect, public policy based on rigorous scientific research, with complexity science as a multidisciplinary bridge, is a first step towards a new way of addressing the major issues of this century.

Authors' contributions
Both authors conceived the study, participated in the design of it, the analysis of the data, and drafted the manuscript. All authors read and approved the final manuscript.

Authors' information
Issa Luna-Pla has a PhD in Media Law with relevant experience on anticorruption policy making. José R. Nicolás-Carlock has a PhD in Physics with relevant training on complexity science research. Both authors are leading researchers of the Observatory of Corruption and Impunity, hosted by the Institute of Legal Research at the National Autonomous University of Mexico.

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

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

References
Altshuler Y, Pentland A (2018) Social physics and cybercrime. In: New solutions for Cybersecurity, 351–364. MIT Press, Cambridge
Alves LG, Mangioni G, Cingolani I, Rodrigues FA, Panzarasa P, Moreno Y (2019) The nested structural organization of the worldwide trade multi-layer network. Sci Rep 9(1):2866
Alves LG, Ribeiro HV, Rodrigues FA (2018) Crime prediction through urban metrics and statistical learning. Physica A 505:435–443

Andvig, J. C., Fjeldstad, O. H., Weltzien, Å., Amundsen, I., Sissener, T. K., & Søreide, T. (2001). Corruption. A review of contemporary research

Ángel A (2017) Duarte. El priista perfecto (Duarte, the perfect PRI member). Penguin Random House, Mexico

Animal Político (2016). Las empresas fantasma de Veracruz (The phantom companies of Veracruz). Accessed: May 2019. https://www.animalpolitico.com/las-empresas-fantasma-de-veracruz/

Arney C (2010) Predictably irrational: the hidden forces that shape our decisions. Math Comput Educ 44(1):68

Ball P (2003) The physical modelling of human social systems. Complexus 1(4):190–206

Barabási AL (2016) Network science. Cambridge University Press, Cambridge

Barabási AL (2016) Network science. Cambridge University Press

Bar-Yam Y (1997) Dynamics of complex systems, vol 213. Addison-Wesley, Reading, MA

Byers J (2017) The physics of data. Nat Phys 13(8):718

Caldarelli G, Wolf S, Moreno Y (2018) Physics of humans, physics for society. Nat Phys 14(9):870

Campedelli GM, Cruickshank I, Carley KM (2019) A complex networks approach to find latent clusters of terrorist groups. Appl Network Science 4(1):59

Capraro V, Perc M (2018) Grand challenges in social physics: in pursuit of moral behavior. Front Phys 6:107

Colliri T, Zhao L (2019) Analyzing the bills-voting dynamics and predicting corruption-convictions among Brazilian congressmen through temporal networks. Sci Rep 9(1):1–11

Colonnelli, E., Gallego, J. A., & Prem, M. (2019). What predicts corruption? Available at SSRN 3330651

Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertesz J et al (2012) Manifesto of computational social science. Eur Phys J Special Topics 214(1):325–346

David-Barrett E, Fazekas M, Hellmann O, Mark L, McCorley C (2018) Controlling corruption in development aid: new evidence from contract-level data

De Domenico M, Brockmann D, Camargo C, Gershenson C, Goldsmith D, Jeschonnek S, Sayama H (2019) Complexity Explained

de Jeude JVL, Aste T, Caldarelli G (2019) The multilayer structure of corporate networks. New J Phys 21(2):025002

DellaPosta D (2017) Network closure and integration in the mid-20th century American mafia. Soc Networks 51:148–157

D'Orsogna MR, Perc M (2015) Statistical physics of crime: a review. Phys Life Rev 12:1–21

Drugov M, Hamman J, Serra D (2014) Intermediaries in corruption: an experiment. Exp Econ 17(1):78–99

Duijn PA, Kashirin V, Sloot PM (2014) The relative ineffectiveness of criminal network disruption. Sci Rep 4:4238

Espinal-Enríquez J, Larralde H (2015) Analysis of Mexico's narco-war network (2007–2011). PLoS One 10(5):e0126503

Fazekas M, Kocsis G (2017) Uncovering high-level corruption: cross-national objective corruption risk indicators using public procurement data. Br J Polit Sci 50(1):155–164

Fazekas M, Skuhrovec J, Wachs J (2017) Corruption, government turnover, and public contracting market structure–insights using network analysis and objective corruption proxies

Fazekas M, Tóth B (2017) Proxy indicators for the corrupt misuse of corporations. October2017:6. U4 -Chr. Michelsen Institute, Bergen

Fazekas M, Tóth IJ, King LP (2016) An objective corruption risk index using public procurement data. Eur J Crim Policy Res 22(3):369–397

Geyer R, Cairney P (eds) (2015) Handbook on complexity and public policy. Edward Elgar Publishing, Cheltenham

Heidenheimer AJ, Johnston M (2011) Political corruption: concepts and contexts (Vol. 1). Transaction Publishers, New Brunswick

Helbing D (2013) Globally networked risks and how to respond. Nature 497(7447):51

Helbing D, Brockmann D, Chadefaux T, Donnay K, Blanke U, Woolley-Meza O, Moussaid M, Johansson A, Krause J, Schutte S, Perc M (2015) Saving human lives: what complexity science and information systems can contribute. J Stat Phys 158(3): 735–781

Hill L (2012) 6. Ideas of corruption in the eighteenth century: the competing conceptions of Adam Ferguson and Adam smith. Corruption 97–112

Holme P, Liljeros F (2015) Mechanistic models in computational social science. Front Phys 3:78

Huntington SP (1970) Political order in changing societies. VRÜ Verfassung und Recht in Übersee 3(2):257–261

Khan MH (2000) In: Jomo KS (ed) Rents, rent-seeking and economic development: theory and evidence in Asia. Cambridge University Press, Cambridge

Lagi M, Bar-Yam Y, Bertrand KZ, Bar-Yam Y (2015) Accurate market price formation model with both supply-demand and trend-following for global food prices providing policy recommendations. Proc Natl Acad Sci 112(45):E6119–E6128

Lambsdorff JG (2007) Causes and consequences of corruption: what do we know from a cross-section of countries? International handbook on the economics of corruption

Livemore MA, Rockmore DN (2019) Law as Data. Computation, Text and the Future of Legal Analysis. The Santa Fe Institute Press, Santa Fe

López-Iturriaga FJ, Sanz IP (2018) Predicting public corruption with neural networks: an analysis of spanish provinces. Soc Indic Res 140(3):975–998

Magliocca NR, McSweeney K, Sesnie SE, Tellman E, Devine JA, Nielsen EA et al (2019) Modeling cocaine traffickers and counterdrug interdiction forces as a complex adaptive system. Proc Natl Acad Sci 116(16):7784–7792

Marshak CZ, Rombach MP, Bertozzi AL, D'Orsogna MR (2016) Growth and containment of a hierarchical criminal network. Phys Rev E 93(2):022308

Méndez F, Sepúlveda F (2009) What do we talk about when we talk about corruption? J Law Econ Org 26(3):493–514

Milinski M (2017) Economics: Corruption made visible. Nat Hum Behav 1(7):0144

Miller JH, Page SE (2009) Complex adaptive systems: an introduction to computational models of social life (Vol. 17). Princeton university press

Mitchell M (2009) Complexity: a guided tour. Oxford University Press, Oxford

Morris SD (2018) Variations on a theme: corruption in Mexico and the US. Can Soc Sci 14(12):13–25

Morselli C, Boivin R (2017) Introduction to the special issue on crime and networks

Mulgan R (2012) Aristotle on legality and corruption. Corruption 25–36

Mungiu-Pippidi A (2017) The time has come for evidence-based anticorruption. Nat Hum Behav 1(0011):1

Mungiu-Pippidi A, Dadašov R (2017) When do anticorruption laws matter? The evidence on public integrity enabling contexts. Crime Law Soc Chang 68(4):387–402

Muthukrishna M, Francois P, Pourahmadi S, Henrich J (2017) Corrupting cooperation and how anti-corruption strategies may backfire. Nat Hum Behav 1(7):0138

Newman M (2018) Networks. Oxford University Press, Oxford

Niu X, Elsisy A, Derzsy N, Szymanski BK (2019) Dynamics of crime activities in the network of city community areas. Appl Netw Sci 4:127

Nye JS (1967) Corruption and political development: a cost-benefit analysis. Am Pol Sci Rev 61(2):417–427

Olken BA, Pande R (2012) Corruption in developing countries. Annu Rev Econ 4(1):479–509

Olmos R (2018) Gigante de lodo. Odebrecht y su historia de corrupción en México (mud giant. Odebrecht and its story of corruption in Mexico). Penguin Random House, Mexico

Ouellet M, Bouchard M, Charette Y (2019) One gang dies, another gains? The network dynamics of criminal group persistence. Criminology 57(1):5–33

Pentland A (2015) Social physics: how social networks can make us smarter. Penguin

Platteau JP (1994) Behind the market stage where real societies exist-part l: the role of public and private order institutions. J Dev Stud 30(3):533–577

Ren XL, Gleinig N, Helbing D, Antulov-Fantulin N (2019) Generalized network dismantling. Proc Natl Acad Sci 116(14):6554–6559

Reno WS (1995) Corruption and state politics in Sierra Leone. In: Corruption and state politics in Sierra Leone. Cambridge University press; African studies series, p 83, Cambridge

Ribeiro HV, Alves LG, Martins AF, Lenzi EK, Perc M (2018) The dynamical structure of political corruption networks. Journal of Complex Networks 6(6):989–1003

Riccardi M, Sarno F (2014) Corruption. Encyclopedia of Criminology and Criminal Justice:630–641

Roldán N, Castillo M, Ureste M (2018) La Estafa Maestra. Graduados en desaparecer el dinero público (the master fraud. Graduates in disappearing public money). Planeta, México

Rose-Ackerman S, Palifka BJ (1999) Corruption and government: causes, consequences, and reform. Cambridge university press, Cambridge

Rumi SK, Deng K, Salim FD (2018) Crime event prediction with dynamic features. EPJ Data Science 7(1):43

Rutherford A, Lupu Y, Cebrian M, Rahwan I, LeVeck BL, Garcia-Herranz M (2018) Inferring mechanisms for global constitutional progress. Nat Hum Behav 2(8):592

Sah S, Loewenstein G (2014) Nothing to declare: mandatory and voluntary disclosure leads advisors to avoid conflicts of interest. Psychol Sci 25(2):575–584

Sayama H (2015) Introduction to the modeling and analysis of complex systems. Open SUNY Textbooks, New York

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504

Smilov D (2010) Anticorruption agencies: expressive, constructivist and strategic uses. Crime Law Soc Chang 53(1):67–77

Sparrow MK (1991) The application of network analysis to criminal intelligence: an assessment of the prospects. Soc Networks 13(3):251–274

Svensson J (2005) Eight questions about corruption. J Econ Perspect 19(3):19–42

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

Transparency Internacional (TI) (2018). The cost of corruption: https://www.transparency.org/what-is-corruption [web, January 16, 2019]

UN General Assembly (2003). Report of the United Nations Convention Against Corruption, October 31, 2003, A/58/422

Varraich A (2014) Corruption: An umbrella concept. QoG Working Paper Series 5(5):3–27

Venturini T, Jacomy M, Jensen P (2019) What do we see when we look at networks. arXiv preprint arXiv:1905.02202

Vespignani A (2012) Modelling dynamical processes in complex socio-technical systems. Nat Phys 8(1):32

Wachs J, Kertész J (2019) A network approach to cartel detection in public auction markets. Sci Rep 9:10818

Wachs J, Yasseri T, Lengyel B, Kertész J (2019) Social capital predicts corruption risk in towns. R Soc Open Sci 6(4):182103

Wandelt S, Sun X, Feng D, Zanin M, Havlin S (2018) A comparative analysis of approaches to network-dismantling. Sci Rep 8(1):13513

Wiesner K, Birdi A, Eliassi-Rad T, Farrell H, Garcia D, Lewandowsky S et al (2018) Stability of democracies: a complex systems perspective. Eur J Phys 40(1):014002

Yoo SH (2008) Petty corruption. Economic Theory 37(2):267–280

Publisher's Note

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