Political Connections and Firms: Network Dimensions

Maurizio Bussolo[a], Simon Commander[b] and Stavros Poupakis[c]

[a] World Bank, 1818 H Street NW, Washington, DC, USA; email: mbussolo@worldbank.org

[b] IE Business School, Maria de Molina 31, Madrid, Spain; Altura Partners, London; and IZA; email: scommander@alturapartners.org

[c] Institute for Global Health, University College London, 30 Guilford Street, London, UK; email: s.poupakis@ucl.ac.uk

(Accepted Version)

Abstract

Business and politician interaction is commonplace. Most research has classified firms as either connected to a politician or not: a binary approach. Yet, there are almost always strong network dimensions to these connections. This paper builds a unique data set to document a network of connections between politically exposed persons, political parties, and firms in seven economies. With this novel dataset, the paper examines the association between the characteristics and performance of firms and the firms' connections with politically exposed person, taking into account the network nature of these connections. The originality of our analysis is to identify how participation and location in a network, including the extent of links, as well as having a strategic location or centrality, is correlated with firm scale and performance. In a binary approach, such network characteristics are omitted. One consequence is that the intensity and consequences of politically connected business may be significantly mis/under-estimated.

JEL: L14, L53, P26.

This is a pre-copyedited, author-produced version of an article accepted for publication in Oxford Economic Papers following peer review. The version of record Maurizio Bussolo, Simon Commander, Stavros Poupakis, Political connections and firms: network dimensions, Oxford Economic Papers, 2022, gpac017 is available online at: https://doi.org/10.1093/oep/gpac017 .

1 Introduction

The ubiquity of networks in the social and economic life of humans is well established.[1] Networks consist of relationships represented as links (or edges) between agents (or nodes). Networks tend to influence behaviour or reflect specific opportunities among the connected individuals, companies and other entities, such as political parties. This paper studies the role of network features – including centrality, clustering and component (big island) – in shaping firm-politician connections as well as firm outcomes.

Several factors motivate adopting a network approach when investigating the political connections of firms. A network approach allows capturing the significant heterogeneity that would otherwise be lost in a binary approach. Further, networks are themselves the mechanism that facilitate privileged access to resources and can act as conduits for the diffusion of information allowing some firms to gain a competitive – and, at times, unfair – edge. They can also provide a resilient framework for political privileges that may prove difficult to dislodge. Public interventions aimed at establishing a level playing field may thus fail if they simply focus on one-to-one relationships.

A useful illustration of the advantages of a network approach is to consider forms of government where power and resources are highly concentrated. In such instances, networks tend to have a hub and spoke appearance with many links going and starting from central nodes with fewer links between nodes further away from the centre.[2] In such cases, a firm connected with a person at or near the hub is likely to be able to extract more rents and monetize the connection(s) to a larger degree than a firm connected with a peripheral node, let alone to a firm that is not connected at all. Consequently, to gain access to the network's political capital, firms may gravitate to politically exposed persons (PEPs), particularly those who have a central location in the network. Fisman (2001) and Ferguson and Voth (2008)

[1] See for example, among sociologists, Mitchell (1969) and Raub and Weesie (1990), and among economists Bala and Goyal (2000), Goyal et al. (2006), and Jackson (2009).

[2] For a discussion of differences in network structure across political systems, see Commander and Poupakis (2020).

have shown that in the context of autocratic regimes, connections can deliver substantial rents to some firms which may be reflected in their equity prices (Rijkers, Baghdadi and Raballand 2015). Yet, these papers do not explicitly consider the network dimensions of connections. Another study by González and Prem (2019) uses a network approach to demonstrate how connected firms in Chile were able to maintain some of their privileges after the transition to democracy. They distinguish between firms in which a director had a direct link with Pinochet and firms with an indirect link. Even if both groups of firms experienced some negative consequences at the end of the Pinochet regime, including a fall in their equity values, there was significant heterogeneity. This points to the importance of adopting a network approach when studying the impact of political privileges.

The objective of gaining access to political networks and the methods used have been found to vary widely but frequently encompass political contributions and lobbying (Jayachandran 2006, Blanes i Vidal, Draca and Fons-Rosen 2012). The returns to connections have also been found to vary substantially, suggesting that connections are used for a wide variety of purposes. Connections can be used to cover behaviour that is detrimental to the firm or to support loss-makers. This latter effect might be expected in contexts where market-based pressures are either weak or where there has been a history of 'soft budget constraints', as in the economies studied in this paper. Networking can also be a strategy for sharing information[3] and for gaining some form of competitive advantage.[4] In the specific context of investment decisions, Fracassi (2017) has shown how social, educational and professional connections among top executives of US public companies, as well as more central locations in social networks, had led them to make similar investment choices that perform better.

Networks also tend to be resilient. Examples include networks tying former SOE man-

[3] Fafchamps and Quinn (2018) cite a variety of papers.

[4] Available evidence indicates that network size and structure vary significantly. For example, dense networks tend to generate bonding capital that can be useful in leveraging valuable assets or connections and hence in ensuring cooperation. By contrast, more diffuse networks with fewer clusters tend to be more suited to providing information or access to information; bridging capital, in other words (bonding and bridging capital are terms employed by Putnam (2000); see also Granovetter (1973).

agers – subsequently owners – to politicians, as well as to others of their type, in Russia and other countries of the Former Soviet Union. In France, research has also shown the power and resilience of networks formed around a common education and professional experience (Bertrand et al. 2018). Indeed, the available evidence suggests that networks can facilitate the transfer of advantages across policy regimes and changes in government.

In this paper, we build a unique and very detailed dataset that identifies PEPs and the networks of connections among them, along with the links between PEPs and firms. The data cover seven economies: Bulgaria, Hungary, Romania, the Russian Federation, Serbia, and the Slovak Republic, as well as a Western European comparator, Spain and allow identifying connections and the features of the networks. With this information, we are able to go beyond the connected versus non-connected (i.e., binary approach) that has been common in the literature. The networks' features that we consider include the number of connections, whether there is a connection with an elected politician, along with the degree to which a person or company is centrally located in a network, as well as whether they are contained within the largest component of the network – the big island – or in a more peripheral part of the network.

It is important to note that our analysis is descriptive, as we cannot claim causality – *from* the connection with a PEP (direct or via a network) *to* firms' outcomes – with the current dataset and analytical approach. The main issue that remains unresolved is that of reverse causality. Connections provide advantages to both sides, firms and politicians. Assuming that firms seek connections to gain privileged treatment – for example, easier access to credit, assets, and infrastructure (Khwaja and Mian 2005, Diwan et al. 2015, Bussolo et al. 2021), or pay lower taxes (Rijkers, Baghdadi and Raballand 2015) – and that these connections affect firms' outcomes may just be part of the story. It is also plausible that politicians look for favours from firms, including by creating employment at opportune moments in the electoral cycle. In this search for votes, politicians may focus their attention to larger firms. Clearly size of the firms and political connections will be correlated in this case, but there is no causation from political connection to size. A brief discussion of the possible sources of exogenous variation in the connections and promising directions of future research is included in the final section of the paper.

While our analysis is descriptive, it still provides valuable insights through its focus on the network dimension. We compare baseline estimates that apply a binary approach indicating whether or not a firm is connected and network estimates, which augment the baseline ones by explicitly taking account of the nature of those connections and their network properties. These two sets of estimates display significant differences suggesting that relying only on a binary approach can be misleading, not necessarily because of any bias in coefficients but by missing heterogeneity. We find that connected firms perform worse but are larger and tend to pay higher wages. Connected firms have important similarities with SOEs who can be considered to exemplify a maximal form of connectedness.

The paper is organized as follows. Section 2 describes the building of the network dataset. Section 3 examines the network features of the political connections in the countries covered by this paper. Section 4 then analyses the association between being connected and firms' outcomes. Section 5 concludes.

2 Description of dataset

We combine two main data sources to build the network dataset employed in our analysis. The first is a database – hereafter referred to as PEPData – that has been assembled using data from major commercial providers of business intelligence (such as Thomson Reuters, Dow Jones) that draw on publicly available information.[5] Our database contains, for each country, an exhaustive list of Politically Exposed Persons (PEP), their associates, such as relatives and business partners, political parties, State Owned Entreprises (SOEs), and private firms. These lists include the names and surnames of the persons, the names of the

[5]Commercial business intelligence providers are collecting this information to reduce business and reputational risk for investors and lenders.

firms, and some basic additional information, such as whether the person is a minister or has some other title or affiliation and, in some instances, the sector of operation of the firms. However, the most interesting feature of these datasets are the links among PEPs, associates, public, and private firms. It is important to highlight that the networks in the PEPData are not made up of connections among elements of the same nature. A social network, for example, comprises all possible 'friendship' among individuals in that network. In the case of the PEPData, people and firms (and political parties) are elements of the network, and the main focus is the link between these two quite distinct elements of the network, rather than the firm-to-firm or person-to-person connections. PEPData provides a snapshot of the networks at a single time point in 2017. These networks build gradually, and the datasets represent them more as a stock variable, the accumulation of connections up to a certain point in time, rather than a flow variable, or the existing connections at a specific point in time. In any case, while the networks may evolve slowly, one cannot assume that they remain constant overtime.

The second data source is the Orbis firms dataset of the Bureau van Dijk. This provides balance sheet and financial information for companies in each of the countries. The matching of the Orbis firms with the firms from the PEPData creates a network dataset which contains information about *both* the connections *and* characteristics and performance of firms.

In the PEPData, a PEP is defined as an individual elected, or appointed, to a politically exposed position, as well as their close relations and associates, whether people or businesses. The PEP definition thus encompasses individuals who are, or have been, entrusted with prominent public functions in a range of institutions, including governments, organisations, and state-owned enterprises. This includes leaders of political parties, members of parliament, senior government officials including both judiciary and legislature, senior military officers, ambassadors, heads of SOEs, and government agencies. Note that this definition of a PEP also includes holding companies owned by individuals who are PEP themselves.

This PEP definition is consistent with widely used definitions, such as that provided by the Financial Action Task Force (FATF)[6], but it also extends the definition to incorporate the business relationships of PEPs, such as interests in firms. Detailed information for each PEP has been assembled from a range of sources that include sanctions, regulatory and legal lists but also a wide variety of media sources. Thus, for each identified individual, we document links to other PEPs, organisations, and firms. By building on these links, we are able to construct national networks of relationships between PEPs and firms. However, despite the common methodology across countries, we expect that relationships between PEPs, and even more between PEPs and firms, will be under-reported in countries with weak institutions and less transparency about political and business ties. This is clearly a potential source of bias and a limitation of our study.

A correct matching between the PEPData and Orbis data is key to building a network dataset useful for our analyses. We employ two methods to match the two datasets. The first method involves taking a firm that appears in PEPData and linking it directly through its name to firm-specific financials contained in the Orbis dataset. Since firm names are used to match the two datasets, we refer to this as the Firm Name Method (FN). In addition, we take information on PEPs from PEPData and match them with firms in Orbis based on whether their names appear in the list of shareholders in Orbis' firms. Because of commonly used names and the danger of mis-measurement, each possible match was subsequently reviewed manually to ensure the integrity of the match.[7] As a result, a substantial number of matches were discarded as false positives, and a revised set of matches adopted.[8] We refer to this method as the PEP Name Method (PN). We then combine the two methods to create our network dataset. This process is described in Figure 1. Finally, we treat all remaining firms

[6]See `https://www.fatf-gafi.org/documents/documents/peps-r12-r22.html` last accessed on 7 March 2022.

[7]Researchers based in each country and with extensive local knowledge, checked each match using a range of complementary sources and documentation. Only verified matches were maintained and false positives and ambiguous matches were discarded.

[8]We explored alternative approaches, such as stochastic matching through location used by Koren et al. (2015) but found that data gaps were too large.

in Orbis as non-connected.

2.1 Descriptive statistics

Table 1 provides basic descriptive statistics, including the number of PEPs identified through each method for each country. Using the FN Method gives significant variation in the number of companies associated with PEPs. For example, in Bulgaria, Hungary, and Serbia only 46-88 firms are identified by this method, as against between 631-851 in Russia and Romania, respectively. The number increases significantly when including the PEP's Name Method (PN). For the consolidated measure (FN+PN), the total number ranges between 384 in Serbia and 4,568 in Russia.[9] Even when using the combined method, the share of connected to total firms in the Orbis dataset remains small.

To get a further sense of scale, we use the assets of PEP-connected companies to yield an approximation of the importance of connected companies in each of these countries. The bottom panel of Table 1 shows that the assets of connected firms, as measured by FN+PN, account for approximately 0.1% to 20% of the total measured assets of firms in Orbis. The ratio is particularly large in Russia and Spain where some massive private firms are connected.[10]

Concerning the properties of connected firms according to the method of identification, applying the size criteria (incorporating revenues, assets, and employment) indicated in Appendix Table A1, it appears that the share of small firms among connected firms is far higher when applying the PN Method; 77% as against 40% for the FN Method. The share of large and extra-large firms comprises 30% when using the FN identification but only 6% when using the PN approach. This pattern is similar in all countries, except Bulgaria. Regarding the size distribution, Spain looks different, as between 27-52% of connected firms are large

[9]Note that for Russia, when using the FN Method, a very large number of firms appeared, attributable to duplication of common firm names (e.g., Sputnik). To avoid this, we use only those firms with a unique name. For the PN Method, each entry was carefully cross-checked to ensure no duplication.

[10]For example, the total assets of just three connected companies amount to over 20% of GDP.

or very large – a far higher share than for the others where the average is between 4-11%.[11] Concerning sector affiliation (see Appendix Table A2), there are differences depending on the method of identification, but they are not that significant. The share of firms in wholesale and retail trade, as well as professional and scientific, is clearly higher for the PN Method, while the reverse is true for financial and insurance services. When comparing connected and non-connected firms, the former is mostly under-represented in manufacturing, construction and trade while being over-represented in all countries in professional and scientific activity.

Table 2 provides mean and median values for a range of firm characteristics and performance indicators distinguishing between connected – as measured by FN+PN – and non-connected firms. These variables come from Orbis and comprise levels of assets, sales, employment, wages, wage per worker, as well as the return on capital employed (ROCE – Earnings Before Interest and Taxes/(Total Assets - Current Liabilities)), and the return on assets (ROA) given as the ratio of net income to total assets.

For several variables, connected firms are consistently larger than non-connected ones. This discrepancy is most salient when the sales outcome is considered, but also for employment and average wages. In the case of ROA and ROCE, connected firms mostly have lower values than non-connected ones. The table also indicates that for both groups of firms, there is a large difference between mean and median values indicating skewness to the right.

3 Network descriptives

By definition, a network is composed of nodes and edges. In our case,[12] the nodes comprise firms (private and SOEs), individuals and political parties and the edges are the links between these entities. It is also important to consider the components of a network. A component comprises all the nodes that are connected to each other (irrespective of their distance) and a network can be divided into different components. The issue of interest is

[11] Country level size and sector breakdowns are available on request.

[12] Firms can also be connected through inter and intra-industry trading, but such transactions (as for intermediate goods and services) are not measured in PEPdata, so this dimension is not considered.

whether a large part of a network falls into one component rather than being composed of small fractional parts. If the size of this large part is much greater than all the others, than this largest component is termed a 'big island'. A network consequently represents a set of relationships, while also providing some form of structure for those relationships.[13]

Figure 2 permits easy visualization of the networks of each country in the respective sub-figure. Each sub-figure shows to the big island of the respective country's network. In these figures, the scaling of each node is done by its degree (number of its connections). Aside from giving a sense of the size and composition of the various networks, the figures bring out some salient country differences. For example, in Russia, political parties are not only less numerous but also less connected to other entities. The network in Russia is heavily influenced by the SOEs, as also private companies, and the links between the two types of firm. By contrast, in Spain, the larger number of political parties stands out, as does the relative absence of SOEs and private firms. Although the scale and location of political parties varies significantly across the other economies, these mappings emphasize the significant place of SOEs – doubtless a legacy of their previous economic systems – in the respective networks. Later, we look at the association between scale and performance and firm-level attributes while explicitly incorporating these network features. Before that, we begin with some description of these networks.

Figure 3 shows that the distribution of firms in terms of the number of PEP connections is quite skewed. The most common group of firms is the one with the smallest number of connections. This skewness parallels that of the size of firms. Further analysis of the type of connections shows that in all countries, although to a lesser extent in both Russia and Spain, the great majority of firm connections for those with only one connection is either with entities or individuals. A trivial share has connections only with other firms. At an economy-wide level, in Russia and Spain, roughly three-quarters of firms have connections with diverse types of entities. In the other countries, this group has a larger share that

[13] A good review of the wider literature on networks is Ward, Stovel and Sacks (2011), see also, inter alia, Goyal, Van Der Leij and Moraga-González (2006), Do, Lee and Nguyen (2015).

reaches between 91-99%. In other words, most connections are not from firms to other firms but involve links from firms to individuals or other entities, such as political parties.

With this in mind, Table 3 concerns only firms that are not exclusively linked with other firms.[14] For them, it is evident that connections with 'individuals' are the most common. To get a sense of how this is calculated, take the case of the 50% share of Russian firms with 2 connections that are linked to individuals. This percentage is calculated as the ratio of the number of firms with two connections that are linked to individuals (515 firms) over the number of firms with two connections (with other firms, individuals, political parties and so on; a total of 1,034 firms).[15] In addition, the larger the number of connections, the higher is the percentage of connections with individuals (at ≥90%). This indicates that when firms have multiple connections, it is almost certain that they will have a connection with an individual (i.e., a PEP). Moreover, firms do not have direct connections with political parties, as the latter connect people rather than firms.

This section has highlighted how individuals (PEPs) are indeed the hubs of the network and, further, that they are more important for larger and more connected, firms. We now deploy a set of network analytics so that we can better understand the properties of these networks.

3.1 Network metrics

Our focus is on measuring, (a) the scale of connections, as indicated by the number of edges a node has, (b) whether links to a specific type of PEP – namely, politicians – matter, (c) the strategic location, notably how much centrality a node possesses, as measured by betweenness and, (d) the locus of where individuals or entities are situated in a network – specifically, whether they are in the largest component or Big Island.

[14] As mentioned in Section 2, our network dataset focusses mainly on the links between firms and individuals, and this is why just few firms have connections exclusively with other firms.

[15] Note that the percentages in Table 3 cannot be summed across rows. Apart from the first column, the rows are not mutually exclusive, as one firm with, say, three connections can appear in multiple rows as it may have one connection with an individual, one connection with a political individual, and one with a party.

We employ measures common in the network literature including, degree – an indicator of scale – and betweenness which represents the extent to which a node falls between other nodes and hence how central particular nodes are in a network. Definitions are as follows. For a node i we have

$$Degree_i = Number\ of\ edges\ connected\ to\ i$$

$$Betweenness_i = \sum_{i \neq j \neq k} \frac{Number\ of\ shortest\ paths\ from\ j\ to\ k,\ through\ i}{Number\ of\ distances\ from\ j\ to\ k}$$

Further, we construct network-level measures;

$$Average\ Degree = \sum_{i=1}^{N} \frac{Degree_i}{N}$$

$$Average\ Distance = \sum_{i=1}^{N} \frac{Avg.\ Distance_i}{N}$$

$$Clustering\ Coefficient = \frac{Number\ of\ closed\ triplets}{Number\ of\ triplets}$$

where a triplet is three nodes having either two edges (i.e., the two nodes connected through the third – open triplet) or three edges (i.e., they are all connected to each other – closed triplet).

On the matter of network size, Table 4 provides the number of nodes in each country, broken down by private firm; state-owned enterprise; political party; political individual and other individuals (i.e., relatives and associates of the political individual). There are significant differences in the size of the network across countries. In absolute terms, Russia has the largest network followed by Spain. Both the Hungarian and Serbian networks are far larger than those of neighbouring Slovak Republic and Bulgaria. Even when adjusted for the size of the population, the network ranking has Russia very clearly at the top followed by Serbia, Hungary, Romania, Slovak Republic and, lastly, Spain.

There are also clear differences in composition: Spain has significantly more political parties and political individuals. In contrast, for Russia not only is there a relatively small number of political parties, but the level and share of SOEs is particularly high. The latter comprise around 9% of the total network, as against an average <3% for the other countries. Romania has a relatively large share of private firms in its network. In all countries, individuals and political individuals comprise between 85-94% of the total network size.

Given our interest in the links to SOEs and private firms, SOE connections are mostly with individuals followed by politicians (the mean share for both is >5%). In Spain, connections to politicians for both SOEs and private firms comprise particularly high shares. In Bulgaria and Russia there is a relatively strong connection of SOEs to other SOEs. For private firms, the picture is much more diverse.

Aside from the absolute size of the network, it is important to look at the components of the network. Table 4 indicates that for all the countries there is indeed a big island that ranges between 31% (Serbia) and 76% (Romania) of the total network size. A larger big island indicates greater integration through the network. Concerning the extent of inclusion of particular types of nodes in the big island, firms and SOEs generally have very high inclusion shares, except for SOEs in Bulgaria. The same is true for politicians with the striking exception of Russia. For political parties, inclusion also tends to be high, except for Spain where only 18% of its (many) parties are in the big island (as against an average of 66% for the other countries). This may reflect the decentralized nature of Spanish politics. None of the countries has any second level component of significant magnitude, reinforcing the point that network activity is concentrated in the big island. Table 4 also shows relatively limited difference in the average degree (or number of links sent to a node) across countries. As regards density, both Bulgaria and the Slovak Republic have a higher ratio of ties in the network relative to the total possible number of ties, while Russia and Spain have the lowest ratios. Pursuing this point, the clustering coefficient further indicates what share of a person/entity's neighbours are neighbours of each other and hence whether dense clusters of nodes are present in the network. Although there is some variation across countries, with higher clustering in Serbia, in general the clustering coefficient is quite low and especially so in Russia. This suggests that most of these networks are not that tightly connected. However, this would not exclude some nodes being concentrated in clusters with an ability to reach other nodes in the population through a small number of bridging connections with low path lengths. This would be characteristic of a small world where the average distance between nodes is very small relative to the size of the population.

Concerning centrality, Table 5 reports the betweenness shares – the extent to which nodes lie between other nodes – for each type of node (the columns sum to 100).[16] The betweenness shares are highest for politicians and political parties, as might be expected. Strikingly, betweenness is also high for SOEs; indeed, in Hungary and the Slovak Republic the share is higher than for either political parties or politicians. Private firms have low betweenness shares across the board.[17] Russia also looks different. The share for both political parties and political individuals is significantly lower than elsewhere, while the share for SOEs is significantly higher. If we extend this analysis to the neighbours of both SOEs and private firms, the betweenness of both SOE and private firm neighbours is particularly high in Russia (77% and 47% respectively).

4 Correlates of being connected

We now examine whether being connected is associated with specific firm level indicators, including sales, wages and return on assets and equity. Our measures of being connected may not be exogenous and unobserved factors, such as the skills of the senior managers of a firm, affecting firm outcomes could also be explanatory factors for having political connections. As

[16] The betweenness share is defined as the ratio of the sum of the betweenness of all nodes in a particular group (parties, SOEs, etc.) over the total network betweenness, i.e. over the sum of the betweenness of all the nodes in the network.

[17] Regarding closeness – which measures how close nodes are to one another and hence of the ability to connect to many, even when not between – shares are higher for both political individuals and other individuals across all countries. For both private firms, and particularly SOEs, the closeness shares are very much higher in Russia than elsewhere.

such, our estimates should be viewed as correlates, rather than indicating causality. Initially, we estimate the following:

$$Outcome_i = \beta_0 + \beta_1 Connected_i + \beta_2 SOE_i + \delta Controls_i + u_i \tag{1}$$

where several outcome measures are related to a dummy variable for whether a firm i is connected, as indicated by the FN+PN method. We also account for whether the firm is a SOE, using information reported directly in Orbis. The reason for including SOEs is that due to their ownership and governance they will, almost by definition, be connected to politicians; features that have, of course, been well documented. The performance measures are ROA and ROCE, as well as the logarithm of sales and average wage. All specifications include the following controls: logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Our analytical sample includes a total of 450,985 observations.

The results from estimating Equation (1) are reported in Table 6 using a pooled country specification. Note that Russia is excluded due to the severe limitations of coverage in Orbis. Consider first the outcomes sales and average wages.[18] These variables do not reflect performance but are more related to the overall scale, or size, at which a firm is operating. In the case of the level of average wages, larger firms pay a higher wage, but it could also reflect the bargaining strength of workers which itself may also be related to the extent of political connections. The coefficient on the dummy for being connected is positive and highly significant for both sales and average wages (columns 1 and 2 in the table), and the same holds for the SOE dummy. Connected firms and SOEs clearly tend to be larger firms. The point estimates indicate that the differences are quite significant. When comparing two private firms that operate in the same sector and country, and have the same other characteristics, the connected firm would have sales 51% larger than those of a non-connected firm and would pay a wage bill that is 26% larger. The 'premium' for sales is even greater

[18] We also used wage bill as an indicator of size and found similar results.

than that for a SOE (18%) while for wages it is similar (28%).

Strikingly, the size premium for connected and state-owned enterprises is not accompanied by a performance premium. Indeed, returns on assets and on capital employed – ROA and ROCE – are lower for connected firms and SOEs vis-à-vis non-connected firms. Once again, as shown by Table 6, the negative coefficients associated with connected firms or SOEs are large and significant. For example, the ROA for a connected firm or an SOE are 2.6-3.1 percentage points lower respectively than the ROA for a non-connected firm. These differentials are equivalent to a drop of 50-60% in the returns of connected firms or SOEs, given that on average the ROA of a non-connected firms is 5.2%.

In sum, our estimates indicate that a measure of being connected is positively associated with levels of activity but that there is a mostly negative and significant association between being connected and firm performance. This pattern is repeated in the case of SOEs.[19] While we believe that the main channel at work is through the benefits that firms extract from connections, this association could have an alternative explanation. Namely, that PEPs might seek out larger firms in the belief that they will be more able to pay rents and in so doing this could weaken performance.

Finally, in looking at whether there are differences in the value of connections due to variation in political regimes across the seven countries, Appendix Table A3 presents the coefficients of SOE and connected separately for each country. While the main results follow for most countries, this is not the case for Serbia (SOE and connected) and Hungary (connected). This might reflect the lag of Serbia's and Hungary's transition process. Indeed, these two were the only countries in the whole EBRD region for which country-level transition indicators have been downgraded. In addition, Serbia ranked much lower than the other countries in our analytical sample, in terms of privatisation and governance and enterprise restructuring (EBRD, 2013).

[19] Note also that a test of equality of the coefficients for connected firms and SOEs, reported in Table 6, is rejected only for sales, and not for the other three outcomes.

4.1 Network regressions

The results reported above use the standard binary approach characteristic of most existing analyses of politically connected business. As such, the emphasis is on the magnitude, sign and significance of the dummy variable for connected as against non-connected firms. However, this approach involves considerable simplification. It ignores the different ways that firms can be connected to political players, let alone the intensity of those connections. As discussed above, location within a network both in terms of being in the big island or not, as well as with respect to centrality, can influence not only the intensity of connections but also how information gets transmitted (Buskens 2002). This mimics in many respects the ways in which, for example, spillovers from foreign direct investment arise (Javorcik 2004). This literature, for example, distinguishes different types of mechanisms and benefits behind the linkages of domestic and foreign firms. In some cases, depending on the country and sector, what matters is being a supplier to a foreign firm, but in others it is being a user of the foreign firm's products or services.

The network approach that we now implement can help address these issues. We focus on the network dimensions in the following sequence. First, we consider simply the number of connections and, second, their nature. Third, we examine information on the network beyond direct connections, i.e., the local network around the firm. Finally, we include network variables that capture aspects of its structure, such as centrality as measured by betweenness, along with belonging to the giant component. The variables that we use in our network specifications are:

(a) Scale: log(Degree+1) which measures the number of connections

(b) Politician: whether the firm has a shareholder who is a politician

(c) Neighbourhood: log(2ndDegree+1) which measures the connections of connections, thereby capturing the possible effect of clustering

(d) Centrality: log(Betweenness+1), conditional on being in the Big Island;

(e) Component: whether the firm is in the Big Island;

A common prior for all these variables is that they will influence the intensity of the connection. This might be through the quantity and quality of information received and/or the extent of privileged access to resources. It will then have an impact on the outcomes associated with being connected. For example, we would expect that being in the largest component of the network confers a stronger advantage than being connected to some other more peripheral component of the network. Similarly, having a larger number of connections will be better than having just a few, while neighbourhood or clustering could also exert an independent influence. Centrality could similarly be expected to confer advantage. However, such advantage might not necessarily show up in better performance if connections are primarily used to facilitate or sanction lower returns on assets or equity, as suggested by the binary estimations reported above.

We now augment Equation 1 by incorporating these network measures, as shown in Equation 2, using the same analytical sample.

$$Outcome_i = \beta_0 + \beta_1 Connected_i + \beta_2 SOE_i + \beta_3 (Network\ Variable)_i + \delta Controls_i + u_i \quad (2)$$

Since the network variables are defined only for connected firms, in all specifications, we replace all network values with zero for all non-connected firms, as well as including a dummy for being connected and for being a SOE. Our objective is to see whether the network variables have explanatory power over and above that for connections alone. In doing this, for each of the indicators: log(Degree+1), Politician, log(2ndDegree+1), log(Betweenness+1), Big Island, we add the network variables separately (Columns 1-5) to the base estimate before including all network variables (Column 6) together.

The main results using these different specifications are shown in Tables 7-10 and, in graphic form, in Figure 4. The tables provide the values for the regressions' coefficients along with their standard errors. The figure represents the results in terms of the estimated dependent variables. The figure highlights the differences in estimated performance between non-connected and connected firms when their connections are handled either on a binary (or baseline) basis or using the network specifications.

Starting with the dummy for being connected, this is consistently positively signed and significant for the scale outcomes, and negatively signed for the performance outcomes, as in the baseline estimation. As expected, the number of connections (log(Degree+1)) also matters, with its coefficient being significant for sales and for both performance variables. In fact, when evaluated at the mean number of connections – which is about 2.6 – this simple network specification produces results that are close to the binary approach. As shown in Figure 4, compared to the 5.2% for a non-connected firm, the ROA for a connected firm is equal to 2.6% in the binary or baseline regression and 3.0% in the network regression. Similar results are found for ROCE; 11.5% for non-connected firms, 6.6% and 7.2% respectively. However, the similarity between the binary and network approach holds only when, as mentioned above, the average number of connections is considered. Indeed, we could think of the binary approach as averaging out the heterogeneity of the network. If, instead of the average number of connections, we consider highly connected firms (as given by those at the 90th percentile having four connections), then the values for both ROA and ROCE are different and substantially lower: 1.8% and 5.2%.

We next turn to the nature of the connection and, in particular, to connections that involve either an elected politician or an appointed PEP. The coefficient for an elected politician is of a large magnitude but significant only for the scale outcomes. Indeed, as shown in Figure 4, being connected to a politician is different from being connected with an appointed PEP. While in the former case, sales are much closer to those for a non-connected firm, in the latter case sales are much larger, and even larger than that estimated with the binary approach. Drawing out this distinction highlights again the advantage of applying a network approach rather than a binary one.

The next specifications extend the analysis to include characteristics of the local network, considering the indirect connections that may exist through connected neighbours, such as second degree, centrality (betweenness) and being part of the main component of the network (big Island). There are similarities between binary and network estimates for the average number of connections, but significant heterogeneity at P90 also holds. The variable that stands out in all specifications is the one measuring betweenness which is significant in all cases, and positively (negatively) associated with the scale (performance) outcomes.

Finally, in the full network specification (column 6 in the tables) our central point about heterogeneity relative to the binary approach stands out.[20] For example, the ROA for a firm connected with a politician, with a 90th percentile value for its degree, is about 0.5% compared to 5.2% for a non-connected firm. This amounts to a reduction of about 90% which is substantially more than the 50% reduction found using a binary approach. In other words, knowing the nature of the connection and the position in the network yields not only useful insights but a much more precise, quantitative assessment of the consequences of the connection.[21]

Our results suggest not only a clear association between connections and performance and scale, but also that the importance of connections is related to specific features of the networks. When compared to the baseline estimates reported in Table 6, our network estimates demonstrate that not incorporating network attributes may not actually bias the coefficient of the connected dummy of equation (1), but that it certainly obscures the large heterogeneity that exists once the network variables are included. For these reasons, taking into account the network dimensions of connections is essential. At the same time, our results also indicate how particular network features – such as having a strategic location in the network – exert more significance.

[20] Tests of joint significance of the network variables cannot be rejected viz., for ROA $F = 2.86$, $p = 0.014$; for ROCE $F = 3.10$, $p = 0.008$; for Log Sales $F = 16.71$, $p < 0.001$; for Log Ave Wages $F = 5.43$, $p < 0.001$.

[21] The coefficient for SOEs continues to be positive (negative) and statistically significant for both scale (performance) outcomes and does not differ from the binary approach.

5 Conclusion

Politically connected businesses are neither a rarity nor limited to SOEs. Their incidence cuts across political systems, regions, and levels of development. Our paper has focused on economies in East and Central Europe, Russia and, as a comparator, Spain. Building a new and unique dataset that identifies PEPs and their links to companies, politicians and political parties, we identify the broad scale of the phenomenon in each of these countries. The originality of the paper lies in our ability to complement this identification of links with information about the complex configuration of the networks of connections that exist in each of the countries.

We show that ignoring the network dimension will lead to a potentially misleading view of how connections function, as well as their consequences. In regimes where power is very concentrated, networks tend to link to, and from, these concentrations of power. This highlights in very stark terms the importance of where in a network a firm sits and to whom it is connected. Yet, this is true also in democratic regimes, even if the network structure is less stark and more complex. In other words, it is not just about being connected (the binary approach) but how and to whom. These features, in turn, materially affect the impact of connections.

The principal interest from our analysis concerns how network features influence these associations. We find clear evidence that location in a network, the extent of links and betweenness or centrality is often positively and significantly associated with firm-level indicators for the scale of activity. Network variables are also seen to be significant in an estimation that has performance indicators on the left-hand side. In the latter instance, the sign is mostly negative and significant. This suggests that connections are being used to sanction weak performance; a situation that could be consistent with a variety of channels, whether tunnelling, transfers or other forms of financing. In short, not only are networks likely to be important in shaping how connections arise and propagate (something that we cannot directly measure), but they are also important in shaping the returns to connections.

Evidence from a range of other studies has suggested that networks tend to be very resilient and are hard to disrupt. In the transition economies covered by this paper, even despite major changes in political and economic regimes, many network attributes appear to have survived the transition and persisted. In the pre-transition epoch, connections were essential in ensuring access to resources and finance. The soft budget constraint was intimately connected to the nature and strength of firm managers' connections. The robust, negative association between performance and connections that we find, whether in binary or network estimates and the positive association with scale, points to features of the past that appear to have been carried forward. Given the region's recent past, it is also hardly surprising that SOEs remain prominent, occupying locations in network space that possess relatively high centrality and displaying a strong negative association with performance.

The data in this analysis cannot establish causality, and the findings that connections matter for firms' outcomes could be due to reverse causality, namely that politicians selectively target larger firms with higher wages but not necessarily more profitable to extract electoral or other benefits. Ideally, to establish the direction of causality, one needs a source of exogenous variation on the formation of these connections. Such empirical strategies may involve a panel dimension of these networks, to explore the impact on the 'strength' of the connection that is linked to the electoral cycle, or the change in the network due to unexpected events, such as death, or other sudden events. While networks are intrinsically resilient and thus tend to be persistent, they are not static and exploring their dynamics – their evolution through time – seem a fruitful avenue for future research.

Supplementary material

Supplementary material is available on the OUP website. This consists of Stata codes used to clean the data and replicate the results, and the online appendix. The data used in this article is obtained from different sources, proprietary (Bureau van Dijk's Orbis data, accessible here https://www.bvdinfo.com/en-us/our-products/data/international/orbis) and covered by confidentiality agreements.

Funding

No funding was used for the current study.

Acknowledgements

The authors thank Wenyu Zhu and Esther Bartl for research assistance and Boyko Amarov, Beata Berkovics, Adrian Hernandez, Marta Kahancova, Leo Rollinger, Alina Scherbakova, and Roxana Voicu-Dorobantu for help with the country-level data matching, as also Katya Ushakova. Thanks to Luis-Felipe Lopez-Calva, Ishac Diwan, Bob Rijkers, and Hans Timmer for very helpful comments, as well as seminar participants at the World Bank and Nuffield College, Oxford.

References

Bala, V. and Goyal, S. (2000), 'A noncooperative model of network formation', *Econometrica* **68**, 1181–1229.

Bertrand, M., Kramarz, F., Schoar, A. and Thesmar, D. (2018), 'The cost of political connections', *Review of Finance* **22**(3), 849–876.

Blanes i Vidal, J., Draca, M. and Fons-Rosen, C. (2012), 'Revolving door lobbyists', *American Economic Review* **102**, 3731–3748.

Buskens, V. (2002), *Social networks and trust*, Kluwer Academic Publishers, Norwell, MA.

Bussolo, M., de Nicola, F., Panizza, U. and Varghese, R. (2021), 'Politically connected firms and privileged access to credit: Evidence from central and eastern europe', *European Journal of Political Economy* p. 102073.

Commander, S. and Poupakis, S. (2020), Political networks across the globe, Technical Report DP No. 13103, Institute of Labor Economics (IZA), Bonn, Germany.

Diwan, I., Keefer, P. and Schiffbauer, M. (2015), Pyramid capitalism: political connections, regulation, and firm productivity in Egypt, Technical Report Policy Research Working Paper No. 7354, World Bank, Washington, DC.

Do, Q.-A., Lee, Y. T. and Nguyen, B. D. (2015), Political connections and firm value: Evidence from the regression discontinuity design of close gubernatorial elections, Technical Report Discussion Paper No. DP10526., Centre for Economic Policy Research, London, England.

European Bank for Reconstruction and Development (EBRD) (2013), Transition report 2013: Stuck in transition?, Technical report, EBRD, London, England.

Fafchamps, M. and Quinn, S. (2018), 'Networks and manufacturing firms in Africa: Results from a randomized field experiment', *World Bank Economic Review* **32**, 656–675.

Ferguson, T. and Voth, H.-J. (2008), 'Betting on Hitler - the value of political connections in Nazi Germany', *Quarterly Journal of Economics* **123**, 101–137.

Fisman, R. (2001), 'Estimating the value of political connections', *American Economic Review* **91**, 1095–1102.

Fracassi, C. (2017), 'Corporate finance policies and social networks', *Management Science* **63**, 2420–2438.

González, F. and Prem, M. (2019), Losing your dictator: Firms during political transition, Technical report, Instituto de Economía. Pontificia Universidad Católica de Chile.

Goyal, S., Van Der Leij, M. J. and Moraga-González, J. L. (2006), 'Economics: An emerging small world', *Journal of Political Economy* **114**, 403–412.

Granovetter, M. S. (1973), 'The strength of weak ties', *American Journal of Sociology* **78**, 1360–1380.

Jackson, M. O. (2009), 'Networks and economic behavior', *Annual Review of Economics* **1**, 489–511.

Javorcik, B. (2004), 'Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages', *American Economic Review* **94**, 605–627.

Jayachandran, S. (2006), 'The jeffords effect', *Journal of Law and Economics* **49**, 397–425.

Khwaja, A. I. and Mian, A. (2005), 'Do lenders favor politically connected firms? Rent provision in an emerging financial market', *Quarterly Journal of Economics* **120**, 1371–1411.

Koren, M., Szeidl, A., Szucs, F. and Vedres, B. (2015), 'Political favor exchange in a democracy', *Central European University, Budapest, mimeo* .

Mitchell, J. C. (1969), *Social networks in urban situations: analyses of personal relationships in Central African towns*, Manchester University Press, Manchester.

Putnam, R. D. (2000), *Bowling alone: The collapse and revival of American community*, Simon & Schuster Paperbacks, NY: New York.

Raub, W. and Weesie, J. (1990), 'Reputation and efficiency in social interactions: An example of network effects', *American Journal of Sociology* **96**, 626–654.

Rijkers, B., Baghdadi, L. and Raballand, G. (2015), 'Political connections and tariff evasion evidence from Tunisia', *World Bank Economic Review* **31**, 459–482.

Ward, M. D., Stovel, K. and Sacks, A. (2011), 'Network analysis and political science', *Annual Review of Political Science* **14**, 245–264.

Tables

Table 1: Number of Connected Firms and Size/Assets

	Bulgaria	Serbia	Slovakia	Hungary	Romania	Russia	Spain
Number of Connected Firms							
FN	46	88	136	63	851	631	638
PN	407	296	1,110	612	1,236	3,937	542
FN+PN	453	384	1,246	675	2,087	4,568	1,180
All	669,642	110,432	439,497	551,846	837,779	6,194,392	1,181,296
Total Assets Ratio							
FN/All	0.02%	1.48%	0.26%	0.06%	2.10%	1.07%	3.47%
(FN+PN)/All	0.15%	3.19%	0.63%	0.09%	2.51%	4.11%	20.30%

Note: FN refers to Firm Name Method (linkage through firm names), PN refers to PEP Name Method (linkage based on names appearing as PEP and as shareholders). Calculations of total assets ratios are among the number of firms in the analytical sample. Source: Authors' calculations using PEPData and Orbis.

Table 2: Descriptive Statistics for Connected and Non-Connected

	BG		HU		RO		RS		SK		ES	
	Non	Conn.	Non	Conn.	Non	Conn.	Non	Conn.	Non	Conn.	Non	Conn.
Mean												
Employees	9.4	20.2	11.7	16.6	7.6	32.6	15.9	38.5	15.5	14.4	18.0	1,784.9
Assets	809k	1,148k	1,172k	1,090k	617k	4,379k	2,044k	7,335k	1,563k	1,378k	8,652k	2,890,000k
Sales	453k	631k	988k	868k	580k	6,970k	1,399k	2,236k	1,356k	841k	3,216k	529,000k
Wage	79k	191k	123k	177k	111k	580k	192k	690k	233k	214k	563k	80,200k
AveWage	4k	6k	9k	10k	5k	7k	7k	10k	15k	16k	31k	59k
ROA	8.6	5.1	5.5	6.2	7.5	3.4	2.0	0.6	3.2	2.0	1.6	1.9
ROCE	15.8	14.2	11.4	20.4	20.4	12.0	2.5	5.3	13.8	10.2	6.6	5.7
Sales/Assets	1.48	0.92	1.76	1.17	1.37	0.98	1.46	1.04	1.47	0.98	1.35	0.87
Wage/Assets	0.38	0.34	0.46	0.31	0.46	0.32	0.40	0.32	0.39	0.33	0.45	0.34
Median												
Employees	2	4	2	2	1	2	2	2	1	1	4	18
Assets	49k	158k	35k	57k	26k	145k	106k	307k	69k	140k	405k	8,894k
Sales	19k	47k	49k	47k	17k	51k	68k	105k	56k	52k	273k	3,115k
Wage	8k	28k	12k	11k	9k	37k	18k	44k	20k	25k	92k	622k
AveWage	2k	4k	7k	8k	4k	5k	5k	7k	10k	12k	26k	44k
ROA	4.8	2.6	3.4	2.4	0.9	0.4	1.5	0.7	2.8	1.0	1.2	1.5
ROCE	10.7	8.7	7.9	9.9	14.2	8.6	2.3	3.1	10.6	6.1	4.3	4.1
Sales/Assets	0.80	0.43	1.11	0.63	0.68	0.45	0.95	0.46	0.96	0.44	0.96	0.45
Wage/Assets	0.15	0.15	0.22	0.13	0.19	0.14	0.16	0.12	0.19	0.13	0.25	0.13

Note: All calculations are are based on the firms in the analytical sample. Conn. refers to the connected firms (FN+PN). Values for Assets, Sales, Wage, and AveWage are in 2016 USD. ROA and ROCE are in percent. Source: Authors' calculations using PEPData and Orbis.

Table 3: Share of firms with 'diverse' connections by entity and number of connections

	Connections with:	Each firm's number of connections: 1	2	3	4	>=5	tot
Russia	Corporate	0	74	93	98	88	77
	Individual	39	50	70	73	95	70
	Pol. Individual	23	10	5	7	21	13
	Pol. Party	0	0	0	0	0	0
Spain	Corporate	0	39	33	33	50	34
	Individual	39	74	89	67	97	77
	Pol. Individual	9	22	44	33	43	31
	Pol. Party	0	0	0	33	0	1
Slovakia	Corporate	0	8	17	20	42	22
	Individual	70	92	83	100	100	90
	Pol. Individual	22	24	33	20	28	25
	Pol. Party	0	0	0	0	0	0
Serbia	Corporate	0	42	40	71	28	26
	Individual	69	100	100	100	100	93
	Pol. Individual	25	0	10	0	24	20
	Pol. Party	0	0	0	0	0	0
Bulgaria	Corporate	0	74	40	67	76	59
	Individual	33	63	80	100	76	68
	Pol. Individual	50	42	0	0	52	41
	Pol. Party	0	0	0	0	0	0
Romania	Corporate	0	21	39	47	40	24
	Individual	23	100	96	87	98	72
	Pol. Individual	66	45	48	47	47	53
	Pol. Party	0	0	0	0	0	0
Hungary	Corporate	0	38	81	61	69	56
	Individual	65	98	100	100	100	95
	Pol. Individual	21	11	12	4	34	23
	Pol. Party	0	0	0	0	0	0

Note: Calculations are based on Firm Name (FN) listing. All values are in percent. Source: Authors' calculations using PEPData.

Table 4: Network size and components (by country)

	Bulgaria	Serbia	Slovakia	Hungary	Romania	Russia	Spain
Parties	31	96	25	38	60	38	319
SOEs	218	195	110	349	205	5,660	476
Firms	504	434	1,344	817	2,731	5,802	892
Politicians	1,745	2,264	1,092	2,322	5,244	18,066	9,631
Other Individuals	1,901	5,144	2,130	4,265	7,068	31,445	9,403
Network Size	4,399	8,133	4,701	7,791	15,308	61,011	20,721
Big Island	1,934	2,499	2,620	5,106	11,593	22,016	13,888
Big Island (%)	44%	31%	56%	66%	76%	36%	67%
Second Largest	51	62	27	18	39	51	98
Parties in BI %	84%	36%	76%	50%	72%	63%	18%
SOEs in BI %	27%	50%	63%	85%	90%	84%	64%
Firms in BI %	70%	58%	73%	77%	91%	64%	75%
Politicians in BI %	57%	47%	68%	82%	83%	29%	89%
Individuals in BI %	26%	20%	38%	53%	64%	26%	46%
Average Degree	2.1	2.7	1.9	2.7	2.8	3.0	3.3
Average Distance	6.3	7.6	6.6	5.9	5.3	5.2	5.1
Clustering Coefficient	1.3%	8.6%	1.7%	0.8%	1.7%	0.4%	1.4%

Source: Authors' calculations using PEPData.

Table 5: Betweenness by type and country

	Bulgaria	Serbia	Slovakia	Hungary	Romania	Russia	Spain
Betweenness parties	33%	26%	24%	21%	32%	13%	38%
Betweenness SOEs	23%	18%	32%	40%	15%	52%	15%
Betweenness firms	2%	8%	2%	1%	1%	4%	5%
Betweenness politicians	36%	39%	30%	25%	38%	20%	35%
Betweenness individuals	7%	9%	11%	13%	13%	11%	7%
	100%	100%	100%	100%	100%	100%	100%

Note: Betweenness is calculated only for nodes in the big island, and for all the rest a value zero is assumed. Source: Authors' calculations using PEPData.

Table 6: Baseline Estimates

VARIABLES	(1) logSales	(2) logAvWage	(3) ROA	(4) ROCE
SOE	0.169***	0.248***	-3.117***	-5.847***
	(0.025)	(0.011)	(0.214)	(0.371)
Connected	0.414***	0.229***	-2.626***	-4.941***
	(0.044)	(0.024)	(0.412)	(0.741)
Observations	450,985	450,985	450,985	450,985
R-squared	0.659	0.645	0.101	0.109
Controls	✓	✓	✓	✓
Wald $\beta_{Conn.} = \beta_{SOE}$	23.32	0.51	1.12	1.20
Wald p-value	<0.001	.475	.290	.274
Non-SOE non-Conn. Mean	12.99	9.56	5.22	11.51

Note: Robust standard errors in parentheses. Significance level is indicated as *** p<0.01, ** p<0.05, * p<0.1. Controls include logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Source: Authors' calculations using PEPData and Orbis.

Table 7: Network regressions: Log(Sales)

	(1)	(2)	(3)	(4)	(5)	(6)
SOE	0.169***	0.169***	0.169***	0.168***	0.169***	0.168***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Connected	0.272***	0.702***	0.639***	0.230***	0.348***	0.696***
	(0.071)	(0.064)	(0.075)	(0.049)	(0.087)	(0.103)
log(Degree+1)	0.131**					-0.269***
	(0.057)					(0.075)
Politician		-0.648***				-0.713***
		(0.083)				(0.120)
log(2ndDegree+1)			-0.068***			0.025
			(0.016)			(0.028)
log(Betweenness+1)				0.048***		0.064***
				(0.009)		(0.012)
Big Island					0.082	-0.005
					(0.101)	(0.141)
Observations	450,985	450,985	450,985	450,985	450,985	450,985
R-squared	0.659	0.659	0.659	0.659	0.659	0.659
Controls	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses. Significance level is indicated as *** p<0.01, ** p<0.05, * p<0.1. Controls include logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Source: Authors' calculations using PEPData and Orbis.

Table 8: Network regressions: log(AvWage)

	(1)	(2)	(3)	(4)	(5)	(6)
SOE	0.248***	0.248***	0.248***	0.248***	0.248***	0.248***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Connected	0.185***	0.291***	0.298***	0.167***	0.174***	0.243***
	(0.045)	(0.032)	(0.041)	(0.031)	(0.056)	(0.058)
log(Degree+1)	0.042					-0.071
	(0.037)					(0.065)
Politician		-0.138***				-0.069
		(0.048)				(0.079)
log(2ndDegree+1)			-0.021**			-0.024
			(0.009)			(0.017)
log(Betweenness+1)				0.017***		0.019**
				(0.005)		(0.008)
Big Island					0.070	0.126
					(0.062)	(0.088)
Observations	450,985	450,985	450,985	450,985	450,985	450,985
R-squared	0.645	0.645	0.645	0.645	0.645	0.645
Controls	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses. Significance level is indicated as *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls include logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Source: Authors' calculations using PEPData and Orbis.

Table 9: Network regressions: ROA

	(1)	(2)	(3)	(4)	(5)	(6)
SOE	-3.112***	-3.116***	-3.117***	-3.112***	-3.117***	-3.111***
	(0.214)	(0.214)	(0.214)	(0.214)	(0.214)	(0.214)
Connected	-1.086	-3.064***	-3.347***	-1.873***	-3.375***	-2.689***
	(0.705)	(0.530)	(0.683)	(0.521)	(0.909)	(0.994)
log(Degree+1)	-1.433***					-1.154
	(0.527)					(0.750)
Politician		0.982				-0.126
		(0.834)				(1.159)
log(2ndDegree+1)			0.217			0.107
			(0.167)			(0.295)
log(Betweenness+1)				-0.198**		-0.157
				(0.081)		(0.120)
Big Island					0.945	2.022
					(1.019)	(1.442)
Observations	450,985	450,985	450,985	450,985	450,985	450,985
R-squared	0.101	0.101	0.101	0.101	0.101	0.101
Controls	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses. Significance level is indicated as *** p<0.01, ** p<0.05, * p<0.1. Controls include logarithm of sales, logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Source: Authors' calculations using PEPData and Orbis.

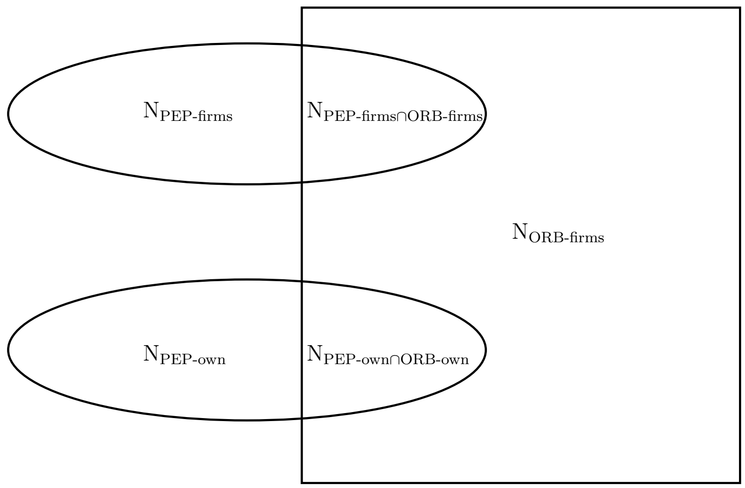
Table 10: Network regressions: ROCE

	(1)	(2)	(3)	(4)	(5)	(6)
SOE	-5.839***	-5.846***	-5.847***	-5.839***	-5.848***	-5.837***
	(0.371)	(0.371)	(0.371)	(0.371)	(0.371)	(0.371)
Connected	-2.098*	-5.581***	-6.039***	-3.463***	-6.311***	-5.099***
	(1.255)	(1.004)	(1.207)	(0.924)	(1.557)	(1.813)
log(Degree+1)	-2.644***					-1.922
	(0.954)					(1.390)
Politician		1.438				-0.346
		(1.481)				(2.075)
log(2ndDegree+1)			0.332			0.085
			(0.278)			(0.492)
log(Betweenness+1)				-0.388***		-0.349
				(0.146)		(0.219)
Big Island					1.730	4.319*
					(1.767)	(2.519)
Observations	450,985	450,985	450,985	450,985	450,985	450,985
R-squared	0.109	0.109	0.109	0.109	0.109	0.109
Controls	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses. Significance level is indicated as *** p<0.01, ** p<0.05, * p<0.1. Controls include logarithm of sales, logarithm of age, logarithm of number of employees, dummy for being multinational, dummy for being exporter, sector dummies, and country dummies. Source: Authors' calculations using PEPData and Orbis.

Figures

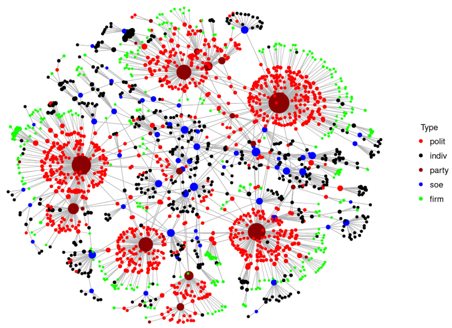
Figure 1: Merging of the two datasets



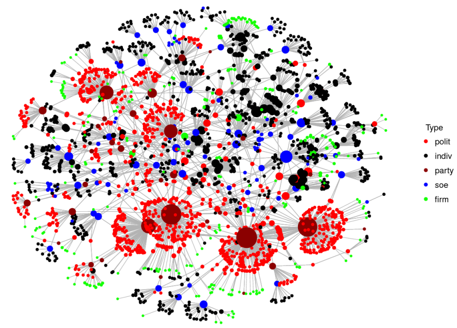
Notes: The ellipse sets represent two lists of the PEPData. One list comprises the names of the firms. The other comprises the list of names of PEPs (politicians and related individuals). The square set represents the ORBIS firm dataset. The total number of names of firms included in the full list of PEPData is equal to the number of firms that are not matched with the ORBIS dataset plus the number of matched firms or: $N_{\text{PEP-firms-Total}} = N_{\text{PEP-firms}} + N_{\text{PEP-firms} \cap \text{ORB-firms}}$ $(12,524 = 10,071 + 2,453)$. Equivalently, the full list of names of PEP individuals connected to a firm in the PEPData is equal to $N_{\text{PEP-own-Total}} = N_{\text{PEP-own}} + N_{\text{PEP-own} \cap \text{ORB-own}}$ $(101,720 = 93,580 + 8,140)$. In the same way, the total number of firms in the ORBIS dataset can be split into three subsets, (a) firms that are matched to the PEPData via the names of the firm, (b) firms that are matched via the name of owner or large shareholder and (c) firms not matched in either way: $N_{\text{ORB-firms-Total}} = N_{\text{PEP-firms} \cap \text{ORB-firms}} + N_{\text{PEP-own} \cap \text{ORB-own}} + N_{\text{ORB-firms}}$ $(9,984,884 = 2,453 + 8,140 + 9,974,291)$.

Figure 2: Network Big Islands

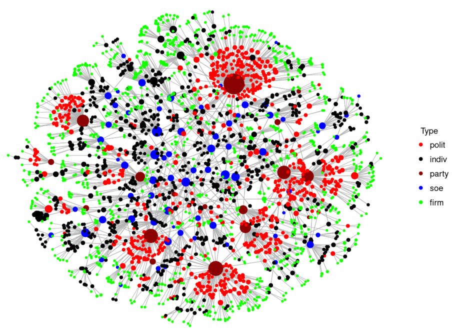
(a) Bulgaria



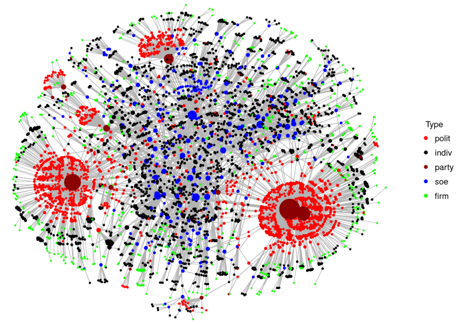
(b) Serbia



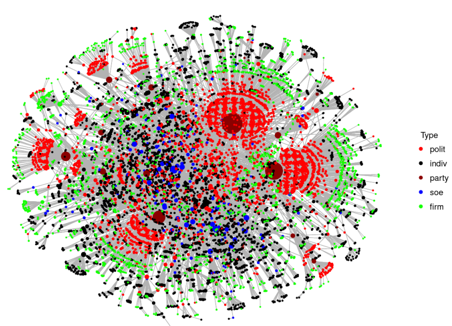
(c) Slovak Republic



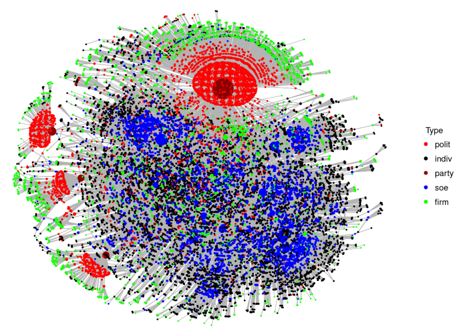
(d) Hungary



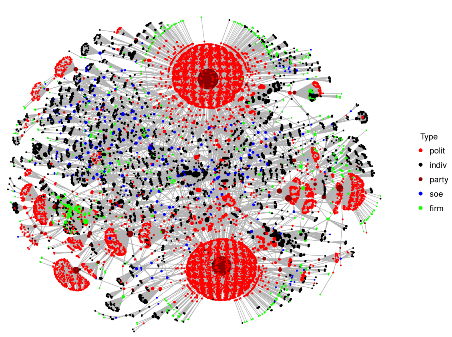
(e) Romania



(f) Russian Federation

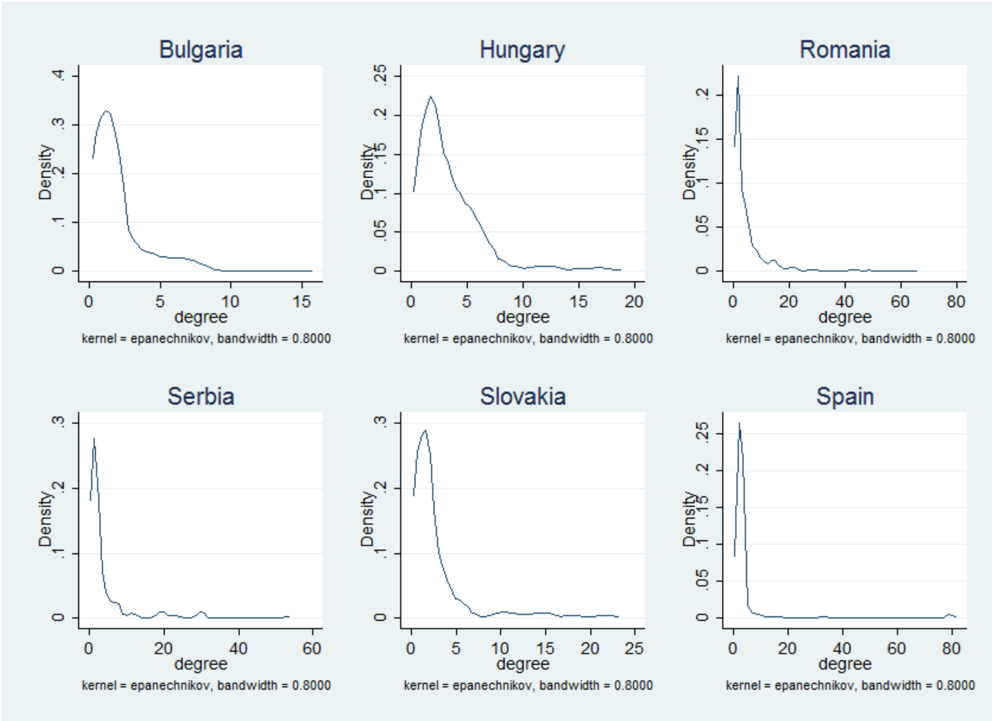


(g) Spain



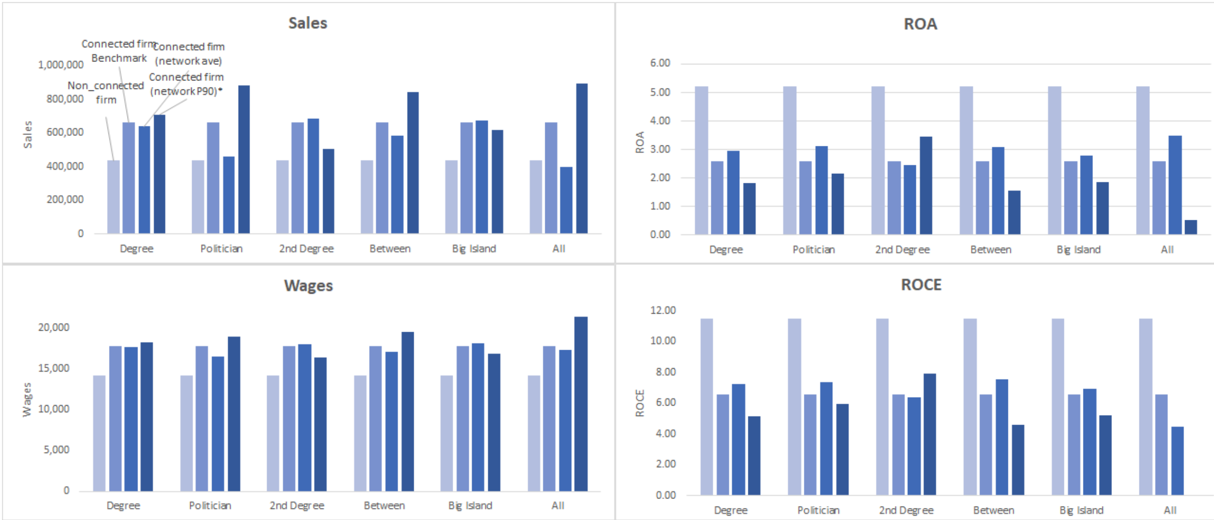
Notes: Note: Graphs of the largest component of each country's network. Size of node determined by its degree. Political parties are coloured in dark red, politicians in red, all other individuals in black, SOEs in blue, and private firms in green. Source: Authors' calculations using PEPData and Orbis.

Figure 3: Firms with only one connection are the largest group



Notes: For each country, the kernel density plot shows the distribution of firms in terms of the number of PEP connections. Source: Authors' calculations using PEPData and Orbis.

Figure 4: Network regressions



Notes: The values of the various bars are calculated as follows: 'Non-connected firm' (lighter blue) is the average level of the variable for this type of firm; the 'Connected-firm Benchmark' uses the coefficient from the binary or benchmark regression and the two network-related bars (darker blue bars) use the regressions coefficients from Tables 7-10 combined with the levels of the respective independent variables. For 'Degree', '2nd Degree', 'Betweenness' specifications, two levels are used; the average level ('network ave') and the level for the 90th percentile ('network P90*'). For the 'Politician' and 'Big Island', the 'network ave' bar is calculated using the coefficient for the dummy for politician and the dummy for being in the Big Island added to the dummy for being connected. The 'network P90*' bar represents firms connected not with a politician nor in the Big Island. Source: Regression results in Tables 6-10.