I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13103

Political Networks across the Globe

Simon Commander
Stavros Poupakis

MARCH 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13103

Political Networks across the Globe

Simon Commander
Altura Partners, IE Business School and IZA

Stavros Poupakis
University of Oxford

MARCH 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Political Networks across the Globe*

Political networks are an important feature of the political and economic landscape of countries. Despite their ubiquity and significance, information on such networks has proven hard to collect due to a pervasive lack of transparency. However, with the advent of big data and artificial intelligence, major financial services institutions are now actively collating publicly available information on politically exposed persons and their networks. In this study, we use one such data set to show how network characteristics vary across political systems. We provide results from more than 150 countries and show how the format of the network tends to reflect the extent of democratisation of each country. We also outline further avenues for research using such data.

JEL Classification: D72, H11

Keywords: political networks, rent-seeking, democratic consolidation

Corresponding author:
Simon Commander
Altura Partners
E-mail: scommander@alturapartners.org

* We would like to thank Laure Bokobza, Maurizio Bussolo, Spyros Kosmidis, Raluca L. Pahontu, and David Weisstanner for helpful comments on previous versions of this draft. Any errors remain ours.

1 Introduction

Power sharing institutions enhance democracy, as they lead to broader inclusion and limit winner-takes-all outcomes. In the seminal work of Norris (2008), such institutions include a proportional representation electoral system, a parliamentary executive, federalism, and established protections of freedom of the press. Although indicative, these characteristics normally overlook informal dynamics. For example, countries in transition may have formally created some of these institutions but other characteristics may still impinge significantly on democratic consolidation. Similarly, regimes that are still classified as autocracies might have experienced recent changes that are not yet fully reflected in the political system. Additional criteria for evaluating political inclusion can therefore be helpful in assessing the quality of democracy or even predict which autocracies are likely to make the transition to democracy.

Social network analysis can be used to improve a comparative approach. Political actors rarely act in isolation and are likely to be influenced by their social connections. This is certainly observed for social ties between policy makers as well as with their connections with the non-political world, such as businesses, friends, and family (Battaglini and Patacchini, 2019). In China, the prevalence of connections is even summarised by the term, guanxi, which, in its worst form can be associated with crony capitalism (Pei, 2016). Although connections can form the basis for corrupt relationships, political connections can also provide easier access to information and reduce the costs of red tape (Bellettini et al., 2013). Consequently, looking at the whole network of connections in the political sphere of a country can have significant explanatory power in a comparative analysis. In particular, the political network can be seen as a reflection of the power structure. In autocracies this usually takes the form of a pyramid under the ruler, followed by the supporting coalition, as well as any tolerated opposition (Boix and Svolik, 2013). By contrast, in democracies political parties play the role of connecting 'hubs' (Miettinen and Poutvaara, 2015).

Comparative analysis has, inter alia, so far studied representations of power dynamics by looking at formal measures, such as the effective number of parties (Laakso and Taagepera,

1979), or even networks of legislature members based on their voting records (Fowler, 2006). However, extending the analysis to broader social networks outside legislatures provides a more holistic approach to the examination of the political sphere. This can reveal informal social ties that can be as powerful as formal institutions.

To date, researchers have assembled data sets on political networks for specific countries that are mostly small scale and relatively limited in coverage.[1] The advent of big data and artificial intelligence has allowed major financial services organisations to collate publicly available information on politically exposed persons (PEPs) and their networks. Although such data have limitations,[2] they provide an opportunity to forge a better understanding of the political networks that exist outside legislatures and other formal institutions. As such, we use a data set on PEPs and their networks from more than 150 countries, compiled by a leading commercial provider of market intelligence in a common and systematic way across all countries. To our knowledge, this is the first empirical analysis of political networks that has global coverage.

The political sphere of a country can be represented by a network, whose members (represented as nodes) include politicians and PEPs, along with their friends and families, political parties, state-owned enterprises (SOEs) and private firms that are affiliated with any of them. Whenever two nodes of this network have a connection with one other, this is represented by edges.[3] Thus, the whole network is a sequence of chains of connections between its members, which can lead to large parts of the network, called components, being grouped together. Like the majority of social networks that exhibit preferential attachment, political networks, too, are characterised by a power law degree distribution, a characteristic of scale-free net-

[1]Some recent examples include, the importance of political networks for electoral success in the Philippines (Cruz et al., 2017), favour exchange (state-owned enterprises' advertisement and media bias) with the media in Hungary (Szeidl and Szucs, 2017), tariff evasion in Tunisia (Rijkers et al., 2015), and firm performance in East Europe (Bussolo et al., 2018).

[2]Such observational data are often incomplete causing estimation problems. See Chandrasekhar and Lewis (2016) for a discussion on measurement error bias.

[3]Such connection (edge) can be either a reported relationship between two politicians or PEPs, or a legal relationship between a PEP and a company, i.e. the former being a shareholder or a member of the latter's board of directors.

works (Watts and Strogatz, 1998; Humphries and Gurney, 2008). This characteristic will naturally lead to the formation of a component that is much larger than the other smaller components, unless this process is disrupted, or even blocked.

In this paper, we explore the size and format of these networks across the globe and their variation by income and political system. We find a strong positive relationship between democracy and the size of the network's largest component. This relationship holds using various measurements of democracy. We find that countries categorised as democracies have, on average, the majority of the political sphere (>50% of nodes) in the largest component. By contrast, the largest component in autocracies accounts for a much smaller share (as low as 30% for the most autocratic regimes). We consider this to reflect the divide-and-rule model proposed by Acemoglu et al. (2004), in which the ruler purposefully interrupts network formation and keeps the political spectrum divided into many smaller components. This way, the ruler does not allow a 'giant component' to emerge, which could lead to their overthrow, and even to democracy.

The paper is organised in the following way. Section 2 formulates the main research hypothesis that we test in this paper. Section 3 describes the procedures followed for the assembly of the network data set, along with a brief description of other country-level data sets that are used in the analysis. Section 4 then lays out the main ways in which political networks vary around the world by providing a visual mapping and narrative. Section 5 looks at the relationship between national networks and the economic and political characteristics of the country. Section 6 concludes by discussing the benefits and limitations of cross-country network analysis and outlines some avenues for future research.

2 Political systems, networks and the giant component

Recent research on network analysis has focused on how individuals within networks are organised, for example their centrality or distance from others, rather than on comparisons

of network-wide measures across countries (Ward et al., 2011). We expand this literature by examining how country-level measures of political networks vary across different political regimes. In particular, we focus on the structure and size of the largest component in each country, often referred to as the 'giant component' (Newman et al., 2002), as it contains a significantly high share of total nodes in the full network. Our exercise is not causal, but mainly descriptive, aiming to explore how these can vary by income and political system. These are evidently important determinants of the "development path" of a society (Lipset, 1959; Acemoglu et al., 2008).

In the political network of a liberal democracy, we should expect nodes in the giant component to be organised around institutions, such as political parties, and, possibly, state-owned-enterprises. Indeed, political parties have been described as "efficiency-improving gate-keepers" who mitigate the effects of networking for rent-seeking (Miettinen and Poutvaara, 2015). By contrast, in autocratic regimes, networks may be centred around powerful individuals, or entities, mainly due to rent-seeking. This type of network organisation has been described as "clan politics" (Collins, 2004) and such organisation can persist even through a process of democratisation (Hale, 2011). This persistence may also have significant consequences for the society, not only with respect to economic performance but also with regard to the distribution of power and income (Bellettini et al., 2013). We examine these aspects graphically and quantitatively.

We hypothesize that the size of the largest component of a country's political network is positively associated with the country's state of democracy. We expect democracies to have a large component that accounts for the majority of the political sphere, whilst in autocracies we expect the largest component, which contains the ruler, to be much smaller and all the rest to be divided among the smaller components. We interpret this as a response to the problem of authoritarian power-sharing, where conflicts emerge mostly among regime insiders, rather than the masses (Svolik, 2012). Indeed, this provides the rationale for divide-and-rule models of personal rule which constrain the connections of the political elite to

ensure that they cannot mount a serious threat to the ruler Acemoglu et al. (2004). Thus, we expect an absence of a 'giant component' in this context. We test this formally with various measures of democracy, accounting for other possible determinants, such as income and possession of natural resources (Robinson et al., 2006; Andersen and Aslaksen, 2013).

3 Data description

3.1 Politically exposed persons

We use a data set of PEPs and their networks, which was made available to us by one of the major worldwide vendors of PEP data. Such data sets are primarily compiled with the aim of money laundering risk mitigation and corruption prevention (Raymond Choo, 2008). We now describe extensively how that data was collected and its principal features.

While there is no global definition of a PEP, most countries and institutions follow the definition of the Financial Action Task Force on Money Laundering.[4] PEPs thus comprise individuals who are currently, or in the last 12 months, entrusted with prominent public functions, including heads of state, ministers and their deputies, members of parliament or legislative bodies, members of political parties' governing boards, members of high-level judicial bodies, such as constitutional courts, diplomats or high-ranking armed forces officers, along with management board members of SOEs. PEPs can be further distinguished between elected and appointed PEPs. In addition, the data also include PEPs close family (parents, spouses and children – although we exclude family members when they are not themselves associated with other individuals or legal entities[5]) as well as any close associates of PEPs (individuals with interests in a legal entity in which the PEP also has an interest, currently or in the past). The systematic collection of information across the globe about these persons

[4]See, for details: https://www.fatf-gafi.org/documents/documents/peps-r12-r22.html.

[5]Simply including all PEPs' family members that have no role in the network will only distort the network measures due to family demographics and transparency. In contrast, when they have other links themselves, they become part of the network and thus their inclusion is important.

and legal entities yields large, predominantly national, networks of nodes represented by persons, political parties, SOEs and private firms, and their respective edges (links) based on the relationship between two nodes.[6]

The data set includes information on individuals drawn from both international and national official records, as well as for individuals, entities and their links drawn from reputable, publicly available media. captured using a combination of Artificial Intelligence (AI) and manual screening by research analysts. It should be noted that measuring PEPs across countries is affected by the very different extents of transparency, access to information and privacy laws that exist. These can either enhance or hinder the identification of PEPs and their affiliates, as well as affecting the overall quality of measurement. In general, although official data should provide reliable evidence, lack of transparency and the like will limit coverage. However, any algorithm for scraping the internet to address such gaps will likely suffer from noise and false positives. The data set that we use mixes both approaches.

The data relate to a single time point, the end of 2017. We do not have any longitudinal data that we could use to study dynamics, such as the evolution of the networks through time. The data set comprises more than 1.2 million nodes (people and legal entities) and more than 2.4 million edges (links between two nodes) for 156 countries.[7] Throughout, we refer to elected PEPs as *Politicians* and to all the other individuals (non-elected PEPs, ex-PEPs, associates, and family members) as *Individuals*. In addition, we group legal entities into state-owned enterprises (*SOEs*), private firms (*Firms*), and political parties (*Parties*).

3.2 Country-level data sets

To implement our analysis looking at the association between network characteristics and features of the wider economy and political system, we also use a set of additional, country-level data sets. Specifically, we use four different sources that measure democracy

[6]Note that all the links in these networks are undirected, i.e. two nodes are connected with an edge and their connection has no direction (for example, one following the other).

[7]We exclude countries that either have a population of less than 300,000 or have missing observations in the main variables we use from the other data resources described below.

and publish annual reports on the state of democracy. In our analysis, we use the 2017 and also the 2007 values of these scores. Although, they are all highly correlated with each other, each has its own advantages and limitations. We now briefly describe each source.

The *Democracy Index* is compiled by the Economist Intelligence Unit. It comprises 60 indicators based on elections, civil liberties, functioning of the government, political participation and culture. These indicators are used to create an index that ranges from 0 (most autocratic) to 10 (most democratic) with variation at two decimal digits (Economist Intelligence Unit, 2018).

The *Freedom in the World* index is published by the Freedom House, a US based non-governmental organisation. The index is constructed based on scores on several categories in two main aspects, civil liberties and political rights. The overall score is an index that ranges from 0 (least free) to 100 (most free) in increments of one. Moreover, countries are categorised based on their scores for the two main aspects into three groups: Not Free, Partly Free, and Free (Freedom House, 2018).

The *Polity Score* is published by the Polity IV project, based at the Center for Systemic Peace in the US. It evaluates democratic and autocratic authority and combines them into a composite index across the spectrum. This is based on six components that capture regulation, competitiveness, openness of executive recruitment, executive constraint, regulation and competitiveness of participation. This results in an integer score from -10 (hereditary monarchy) to 10 (consolidated democracy) (Marshall et al., 2018).

The *Liberal Democracy Index* is compiled by Varieties of Democracy (V-DEM) hosted at the University of Gothenburg. The index aims to captures the state of liberal democracy around the world. It is based on measures of civil liberties, rule of law, independent judiciary, checks of the executive power, and freedoms of association, election, expression, elected officials and suffrage. It ranges from 0 (worse) to 1 (best) and has values at three decimal points (Coppedge et al., 2019).

Finally, we use as a measure of income, GDP per capita PPP, (current international $),

taken from the World Development Indicators (WDI) online data set (World Bank, 2019). We also apply the World Bank classification to group countries into the following categories: Low Income, Low Middle Income, Upper Middle Income and High Income. In addition, we use total natural resource rents as a share of GDP, as high resource rents have been fund to be associated with behaviour that might affect the structure of the networks.

4 Visualising Political Networks

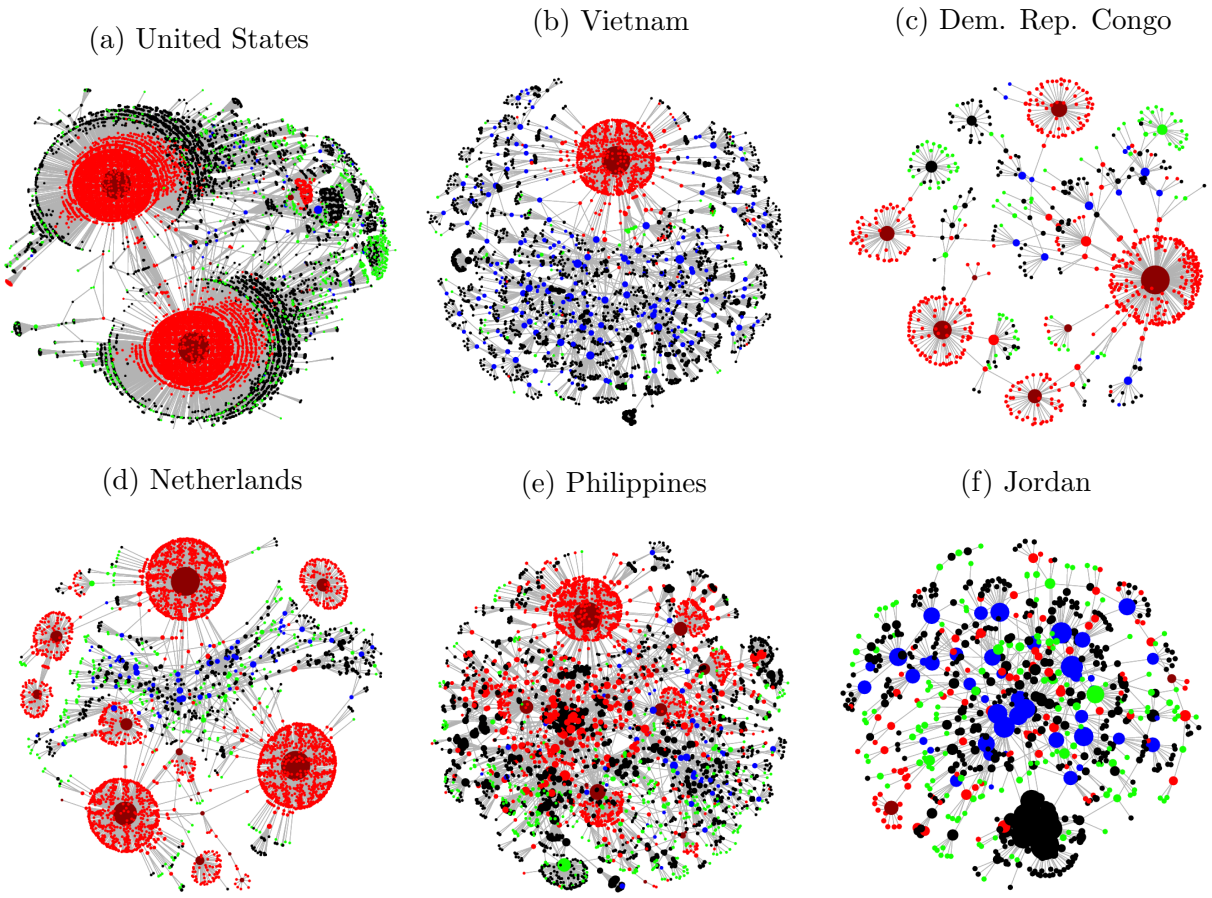
A network not only represents relationships but also provides some form of structure for those relationships. That structure may, in turn, be informative about the sort of opportunities that exist. When measuring networks, the main descriptive elements are nodes - denoting individuals or entities - and edges or links between those nodes. Networks are also commonly represented in terms of degree (the number of links sent to a node) and density (as indicated by the ratio of links in a network to the total possible number of links). In what follows, we apply measures that are able to summarize network structure for each country. We also focus on the place of specific nodes in those networks.

We start with mappings of six selected country networks (Figure 1) that have been selected as being representative of different political systems and income levels.[8] In order to make the figures easier to read, we focus on the largest component (which is defined as a subset of nodes where all its members are connected with at least one other node of the same subset) and scale the size of the nodes by degree. In all figures, politicians are shown in red and all other individuals in black, while political parties are shown in dark red, SOEs in blue, and private firms in green.

The top row presents networks for two high-income consolidated democracies, namely, the United States (1a) and the Netherlands (1d). The political networks of countries of this type are characterised by a concentration of the majority (both around 55%) of the nodes in the largest component forming a big island. These networks are natural representations

[8] A full set of country network maps is available on request.

Figure 1: Examples of country networks



Note: Each network map reflects the largest component of the network of people and legal entities in each country. Size of nodes is based on the number of edges each node has. Political parties are coloured in dark red, politicians in red, all other individuals in black, SOEs in blue, and private firms in green.

of scale-free networks with preferential attachment (where new nodes are more likely to connect with nodes that already have a high number of connections), resulting in power-law degree distributions within these countries. However, there is a feature that separates these countries into two groups, namely, the electoral system. In countries with first-past-the-post electoral systems, we can see the powerful role of the two major parties (the US being the most prominent, but also countries like the UK, Canada, and India). In contrast, in countries with proportional representation, like the Netherlands, several parties form hubs (some being more important than others, i.e. having more nodes connected to them or having a smaller degree of separation to the rest of the network).

In the middle row, two middle income, South East Asian countries with differing political systems are mapped. Vietnam (1b) is an example of a one-party socialist state, albeit one with growing tolerance of a market economy. The largest component includes 40% of the nodes in the network and its structure resembles that of a star network, due to its orientation around the one ruling political party. The Philippines (1e) is a functioning, if flawed, democracy. However, due to its autocratic past there is a somewhat similar network structure. The largest component is also realtively small and accounts for only 29% of nodes. Within this component, although there is more than one political party and the structure is far from a star network, we see a very different picture than in the previously described consolidated democracies. Apart from the political parties, there are many politicians and other individuals with prominent places in the network; doubtless a reflection of the long history of cronyism in the country.

Finally, the bottom row gives examples of low and low-middle income countries in Africa and Middle East. The Democratic Republic of Congo (1c) is an example of a hybrid regime, with a network structure very similar of that in Philippines, with the largest component representing only 24%. However, in this instance, the principal nodes are the political parties and not individuals, while the number and location of the parties reflects the electoral system of proportional representation. In contrast, Jordan (1f) presents an example of another type of hybrid regime, with many features similar to other autocratic states in the region. The largest component only accounts for 33% of the total network and there is an absence of political parties in the network structure. Instead, there are SOEs and clusters (whether one or many) of individuals, not all of whom are politicians. Indeed, this is a characteristic of many other monarchies, reflecting the predominance of a very restricted number of influential people or families, as well as a strong concentration of political power.

These examples already highlight significant cross-country variation both in network size and composition while also suggesting the association between these characteristics and political system and income. The next section explores such variation in greater detail.

5 Quantitative Findings

5.1 Descriptives statistics

We start by presenting some descriptive statistics at a country level. Table 1 presents summary statistics for the main characteristics of the country networks. The bottom panel also includes the additional variables that we use in our analysis and which have been drawn from the various country-level data sets mentioned above. For our analysis, we use information on 153 countries for which we have complete information for all variables.

Table 1: Summary of all variables

Variables	Obs.	Mean	St. Dev.	Min	Max
Nodes	153	7,924	15,002	323	94,023
Edges	153	15,389	30,012	685	212,996
Nodes per capita	153	0.47	0.47	0.02	2.49
Edges per capita	153	0.89	0.89	0.04	4.33
SOEs	153	0.04	0.03	0.01	0.25
Firms	153	0.06	0.05	0.01	0.28
Individuals	153	0.51	0.13	0.19	0.85
Politicians	153	0.39	0.15	0.05	0.77
Share in Largest Component	153	0.46	0.18	0.03	0.77
Democracy: Economist	153	5.54	2.14	1.50	9.87
Democracy: Freedom House	153	57.66	28.76	3	100
Democracy: Polity IV	153	4.37	5.95	-10	10
Democracy: V-Dem	153	0.42	0.26	0.03	0.87
GDP per capita	153	20,840	21,941	738	124,609
Resource rents as % GDP	153	7.70	9.84	0.01	42.67
log(GDP per capita)	153	9.35	1.20	6.60	11.73
log(Resource rents as % GDP)	153	0.78	2.16	-7.76	3.75

Notes: Top panel includes the variables created in the network data set. For the per capita values, the population estimate of 2017 is used from the WDI. The variable share in largest component is based on authors' calculations using the network data set. Bottom panel includes all variables taken from various institutional data sets: WDI, EIU, PolityIV, FH, and V-Dem.

First, we look at the number of nodes and edges. The country with the most nodes and edges is Brazil, followed by China and Mexico, while at the bottom of the list are small countries like Guyana, Gambia, Lesotho, and Surinam. Since these two measures are heavily influenced by the size of a country, we calculate their per capita expressions (using 2017 country population estimates taken from WDI). Expressed in per capita terms, Mexico

and Brazil rank just around the median, while China ends up at the very bottom. The countries that come out at the top are among the smallest in size but with complex political networks. Most are offshore financial locations, such as Malta, Cyprus, and Panama.

Second, we calculate the composition of the network looking at SOEs, Firms, Politicians and Individuals, expressed as shares of the total number of nodes. SOEs are on average the smallest share (4%), with a maximum of 25% found in Cuba, reflecting its state-controlled economy, whereas the US has less than 0.5%. The countries with the largest shares of firms in the network tend to be resource-rich countries like UAE, Qatar, and Azerbaijan, but also tax havens such as Luxembourg, Cyprus, and Malta. Turning to politicians, on average their share is 38%, with the countries at the top being Ethiopia, Myanmar, and the D. R. Congo, while at the bottom are Singapore, Bangladesh and almost all the Gulf countries. In the last row we present the share of nodes belonging to the largest component. The countries with the highest shares for the largest component are Switzerland and Kenya, while at the bottom is Niger with just 3% of the nodes belonging to the largest component.

In the bottom panel of Table 1, we also present summary statistics for the country-level variables that we use later in the regression analysis. We present the four measures of democracy, which are all highly correlated with each other (between 0.80 and 0.95). However, the measures of GDP per capita and Resource Rents are heavily skewed so in the regressions we use their logarithmic transformations (also presented in the Table).

5.2 Network features by Income and Democracy

We now look at whether the variables from the top panel of Table 1 vary across two dimensions: income and democracy. First, we group the countries by their level of income, following the World Bank classification of high, upper middle, lower middle, and low income categories. Second, we group by political system using the classification from Freedom House: not free, partly free, and free.

Table 2 presents the means and standard deviations according to the income classification.

The table shows that network size, as measured by nodes and edges per capita, increases with income. Indeed, there is a very sharp increase from lower to upper middle income. High income countries have on average networks with more than three times as many nodes per capita as low middle income countries and seven times as many as a low-income country. Regarding the composition of the network, the share of firms in high income countries is around double that for other categories. Politicians comprise a significantly higher share in the two lower income categories. Finally, the share of nodes belonging to the largest component increases with income. In particular, starting from 36% for low income countries, it rises to 56% for high income countries. All differences are statistically different at a 1% significance level calculated using a Kruskal-Wallis rank test performed separately for each variable (presented in the last column).

Table 2: Network across income levels

	Low Income	Low Middle Income	Upper Middle Income	High Income	KW χ^2
Nodes	1,571	4,756	14,830	8,116	31.27***
	(696)	(5,632)	(25,294)	(8,716)	
Edges	2,953	8,798	28,141	16,893	35.73***
	(2,611)	(11,350)	(50,309)	(18,847)	
Nodes per capita	0.11	0.24	0.66	0.72	64.51***
	(0.08)	(0.23)	(0.46)	(0.55)	
Edges per capita	0.21	0.41	1.19	1.47	69.99***
	(0.20)	(0.41)	(0.82)	(1.05)	
SOEs	0.02	0.03	0.04	0.04	13.74***
	(0.02)	(0.02)	(0.05)	(0.02)	
Firms	0.04	0.04	0.06	0.09	27.75***
	(0.03)	(0.02)	(0.06)	(0.06)	
Individuals	0.45	0.49	0.52	0.55	13.38***
	(0.13)	(0.12)	(0.12)	(0.13)	
Politicians	0.46	0.43	0.37	0.32	23.41***
	(0.14)	(0.12)	(0.13)	(0.16)	
Largest Component	0.36	0.41	0.46	0.56	23.89***
	(0.20)	(0.15)	(0.20)	(0.12)	
Observations	26	42	42	43	

Notes: Income categories Low, Low Middle, Upper Middle, and High Income are defined as in World Bank's 2017 WDI. KW-test χ^2 refers to the statistic from the Kruskal-Wallis test, with ***p<0.01, **p<0.05, *p<0.1

Table 3 shows how the network variables vary by Freedom House status. Nodes and edges

per capita are about 50% higher for partly free compared to not free, but then double for countries with free countries. While the shares of SOEs, firms, politicians and individuals vary significantly across income categories, this is far less the case for the Freedom House status. Only the share of firms varies significant, but the differences are not substantial and are not linear. As with income, the share of nodes in the largest component increases with the degree of freedom. While countries that are not free or partly free have 36% and 43% of the nodes in the largest component, respectively, free countries have 54% of them. In other words, free (as well as high income) countries have, on average, the majority of nodes in the largest component, whereas for all the groups the majority is, on average, divided across smaller components.

Table 3: Network across political systems

	Not Free	Partly Free	Free	KW χ^2
Nodes	6,296	6,091	10,706	11.26***
	(15,742)	(12,202)	(16,563)	
Edges	12,786	11,029	21,146	13.11***
	(34,652)	(20,097)	(33,586)	
Nodes per capita	0.26	0.36	0.70	33.37***
	(0.34)	(0.36)	(0.53)	
Edges per capita	0.47	0.67	1.37	35.11***
	(0.59)	(0.69)	(1.01)	
SOEs	0.05	0.03	0.03	5.4*
	(0.05)	(0.02)	(0.02)	
Firms	0.07	0.05	0.06	10.18***
	(0.07)	(0.04)	(0.04)	
Individuals	0.47	0.50	0.53	5.3*
	(0.13)	(0.12)	(0.13)	
Politicians	0.40	0.40	0.36	3.62
	(0.16)	(0.13)	(0.14)	
Largest Component	0.36	0.43	0.54	27.32***
	(0.17)	(0.17)	(0.15)	
Observations	40	54	59	

Notes: Freedom status is defined as Free, Partly Free, and Not Free, as this is calculated by the Freedom House. KW-test χ^2 refers to the statistic from the Kruskal-Wallis test, with ***p<0.01, **p<0.05, *p<0.1

5.3 Correlates of the Size of the Largest Component

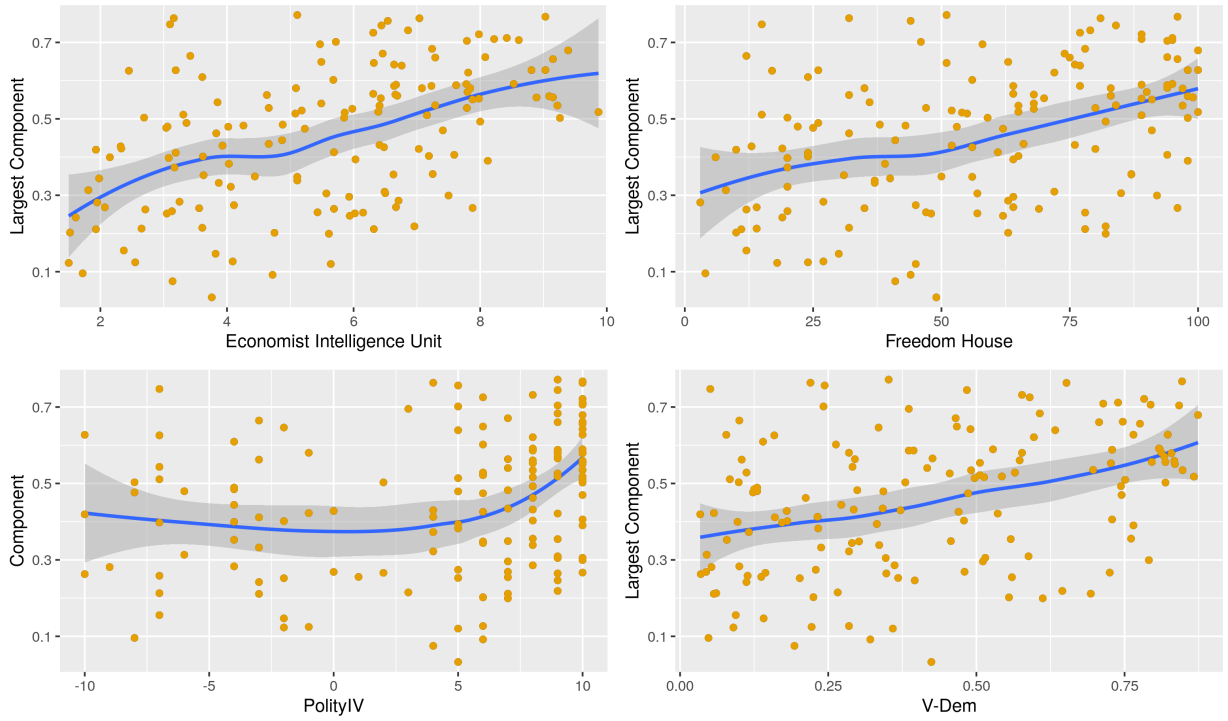
We start by visualising the relationship between democracy and the share of the largest component in a country's network. Figure 2 presents four scatter plots, one for each measure of democracy. On the y-axis is the share of the largest component, while on the x-axis is the democracy index. We also include a fitted line based on a simple local polynomial regression, along with its 95% confidence interval. In other words, we start by regressing the largest component share on the democracy index only.

It can be seen that the share of the largest component rises with the extent of democracy. This is true for all measures of democracy. In particular, for the measures from the EIU, Freedom House, and V-DEM the increase is almost linear. However, when we use the Polity IV, we see that the fitted line is flat for autocracies and anocracies (-10 to +5), whereas for values above +5 (democracies) it increases sharply. Moreover, it is clear from all plots that the expected share of the largest component based on democracy passes the 50% threshold for democracies. In other words, countries that are democracies can be expected to have the majority of the political network contained in the largest component.

We now explore the relationship between the largest component share and democracy through linear regression analysis. In particular, we estimate a series of specifications using the share of the largest component as the dependent variable. We estimate eight models, presented in Table 4. In the first four columns, we include in the right-hand side the various measures of democracy and the logarithm of GDP per capita. In the next four columns, we also include the logarithm of resource rents (as % of GDP), and total number of nodes in the country's network. All the covariates are presented standardised, so the coefficients can be interpreted directly as the change in the share of the largest component for one standard deviation increase in each covariate.

Looking at the first four columns of Table 4 we see that democracy is positively associated with the share in the largest component, irrespective of which index we use. A one standard deviation increase in the democracy index of a country is correlated with a 4-7 percentage

Figure 2: Largest Component correlates with Democracy and Income



Notes: The y-axis contains the share of nodes in each country that belong to the largest component. The x-axis contains the score each country gets in the respective democracy measure. The fitted line is based on a simple local polynomial regression with the shaded area showing the 95% confidence interval.

points increase in the share of the largest component depending on the index and that is always significant at the 1% significance level. Moreover, a one standard deviation increase in the logarithm of GDP per capita is associated with a 3-6 percentage point increase, which is significant at 1% for three of the four specifications. Interestingly, looking at the various measures, we see that in all specifications – apart from the one using Polity IV – the magnitude of democracy is larger than the magnitude for income. In particular, looking at the first specification, which uses the EIU index - the one that has the largest explanatory power (R-squared is 0.26) compared to the other three - we see that the democracy coefficient is more than double the income coefficient.

We then include the logarithm of Natural Resource Rents as a % of GDP as this might affect the baseline relationship. We also include the logarithm of the number of nodes as a control for the size of the networks, and also as a proxy for the size of the countries.

Table 4: Country-level regressions of Share of Largest Component

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std(Economist)	0.072***				0.073***			
	(0.014)				(0.016)			
Std(Freedom House)		0.054***				0.052***		
		(0.014)				(0.016)		
Std(Polity IV)			0.039***				0.034**	
			(0.013)				(0.014)	
Std(V-DEM)				0.048***				0.046***
				(0.014)				(0.015)
Std(log(GDP per capita))	0.028*	0.042***	0.061***	0.044***	0.025	0.035**	0.051***	0.035**
	(0.015)	(0.015)	(0.013)	(0.015)	(0.017)	(0.016)	(0.016)	(0.017)
Std(log(Rents))					0.008	-0.001	-0.006	-0.003
					(0.018)	(0.018)	(0.019)	(0.019)
Std(log(Nodes))					0.017	0.020	0.018	0.022
					(0.016)	(0.016)	(0.016)	(0.016)
Dependent Mean	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457
Observations	153	153	153	153	153	153	153	153
R^2	0.255	0.213	0.187	0.198	0.263	0.223	0.196	0.210
Adjusted R^2	0.245	0.203	0.176	0.187	0.243	0.202	0.174	0.189

Notes: All independent variables are standardised, so the coefficients represent the change in the dependent for one standard deviation increase in the respective variable. Standard errors in parentheses are robust to heteroskedasticity and serial correlation. Significance levels *p<0.1; **p<0.05; ***p<0.01

Larger countries have bigger networks which, as a consequence, might result in a larger share of the largest component. The results of these estimates are shown in columns (5) to (8) of the same table. The conclusions based on the previous results are maintained and are even stronger. While all democracy measures have similar magnitudes and significance, the coefficient of the logarithm of GDP per capita is reduced substantially for all specifications, remaining significant at 1% for only two of the four models. While resource rents and the number of nodes have the expected sign for most specifications - negative and positive, respectively - there is no case in which these are significant at any conventional significance level. Accordingly, we maintain the first four columns as our preferred specification. We also experimented with adding to the specifications, quadratic terms for democracy and income, as well as their interaction, but none were significant.

Finally, as a robustness check, we re-estimate all specifications using the lag of democracy and the lag of the logarithm of GDP per capita. The structure of the current networks is

likely the result of a longer process of democratic consolidation, which extends possibly many years back, rather than just the contemporaneous status of democracy. We replace democracy and income with the same measures as recorded in 2007. The results of these estimations are presented in Table A1. We see that the results hold and are very similar to the main results for all specifications.

6 Conclusion

Our paper has taken a unique data set assembled with a common methodology and global coverage of political networks and mapped the varying structures and composition of networks across a very large group of countries. At this stage, the approach is primarily descriptive as we aim to document individual country networks, while also starting to explore how these networks' composition varies, controlling for factors such as income and political system.

We are able to confirm our main hypothesis that the size of the largest component of a country's political network – an indicator of the extent of integration of a network – is positively associated with the country's level of democracy. We show this by looking at some specific cases, comparing means by categories of democracy, and by performing regression analysis using almost all countries with a population greater than 300,000. In addition, we show that the tipping point of the association between the share of the largest component and democracy is 50%, consistent with our hypothesis that in democracies, the majority will be in the largest component which effectively forms a giant component. This can also be considered an empirical test of the divide-and-rule model proposed by Acemoglu et al. (2004) and the broader conjectures about the way in which power and patronage is organised in societies where institutions are weak.

Several possible avenues for future research emerge from this work. Research on networks has indicated that a central or strategic location in a network is associated with higher

benefits as opposed to being in the periphery (Hojman and Szeidl, 2008). Furthermore, centrality may provide a clue as to how to address networks in the event that policy-makers wish to disrupt or limit the effectiveness of a particular network of individuals and/or firms. Targeting central nodes, for example, could potentially create disruptive effects. In addition, networks might differ according to the ways in which democracies are organised, as for example presidential versus parliamentary systems, or whether presidents or prime ministers have higher centrality in a network. Presidential systems would be expected to have more star shaped networks compared to those with divided executive power, as discussed in (Hale, 2011). Similarly, it would be rewarding to assess whether members of particular types of political parties for example, right or left-leaning parties - have higher centrality, as also whether they are, on average, more connected to private companies or to SOEs. Other aspects that would warrant further exploration would be age and gender both with respect to participation in political networks and their position in those networks as measured by their centrality relative to others.

A further extension could be to look at the location and centrality of persons and entities that have been, or are, the target of sanctions. Since the avowed aim of such sanctions is usually to target specific individuals or organisations, it may be possible to evaluate and even predict the effectiveness of sanctions. This might include a focus on whether sanctions target the core or periphery of a network or whether they target isolated individuals, as against hubs or bridges which would be nodes that have a central role in connecting large parts of the network. If longitudinal data on these networks becomes available, it would be possible to extend our analysis to focus more on how networks develop and change in response to electoral cycles or regime changes.

References

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3):808–42.

Acemoglu, D., Verdier, T., and Robinson, J. A. (2004). Kleptocracy and divide-and-rule: A model of personal rule. *Journal of the European Economic Association*, 2(2-3):162–192.

Andersen, J. J. and Aslaksen, S. (2013). Oil and political survival. *Journal of Development Economics*, 100(1):89–106.

Battaglini, M. and Patacchini, E. (2019). Social networks in policy making. *Annual Review of Economics*, 11:473–494.

Bellettini, G., Ceroni, C. B., and Prarolo, G. (2013). Political persistence and economic growth. *European Journal of Political Economy*, 31:165–179.

Boix, C. and Svolik, M. W. (2013). The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships. *Journal of Politics*, 75(2):300–316.

Bussolo, M., Commander, S., and Poupakis, S. (2018). Political connections and firms: Network dimensions. WP 8428, World Bank Policy Research.

Chandrasekhar, A. G. and Lewis, R. (2016). Econometrics of sampled networks. Stanford University, Mimeo.

Collins, K. (2004). The logic of clan politics: Evidence from the Central Asian trajectories. *World Politics*, 56(2):224–261.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., et al. (2019). *V-Dem Dataset v9*. Varieties of Democracy (V-Dem) Project.

Cruz, C., Labonne, J., and Querubin, P. (2017). Politician family networks and electoral outcomes: Evidence from the Philippines. *American Economic Review*, 107(10):3006–37.

Economist Intelligence Unit (2018). *Democracy Index 2018. Me too? Political participation, protest and democracy*. Economist Intelligence Unit, London.

Fowler, J. H. (2006). Connecting the Congress: A study of cosponsorship networks. *Political Analysis*, 14(4):456–487.

Freedom House (2018). *Freedom in the World 2018: Democracy in crisis*. Freedom House, Washington, DC.

Hale, H. E. (2011). Formal constitutions in informal politics: Institutions and democratization in post-Soviet Eurasia. *World Politics*, 63(4):581–617.

Hojman, D. A. and Szeidl, A. (2008). Core and periphery in networks. *Journal of Economic Theory*, 139(1):295–309.

Humphries, M. D. and Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PloS One*, 3(4).

Laakso, M. and Taagepera, R. (1979). "effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, 12(1):3–27.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53(1):69–105.

Marshall, M. G., Gurr, T. R., and Jaggers, K. (2018). *Polity IV project: Political regime characteristics and transitions, 1800-2018*. Center for Systemic Peace, Vienna, VA.

Miettinen, T. and Poutvaara, P. (2015). Parties as efficiency-improving gatekeepers in rent-seeking societies. *European Journal of Political Economy*, 38:87–101.

Newman, M. E., Watts, D. J., and Strogatz, S. H. (2002). Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572.

Norris, P. (2008). *Driving democracy: Do Power-Sharing Institutions Work*. Cambridge University Press, Cambridge.

Pei, M. (2016). *China's crony capitalism*. Harvard University Press, Cambridge, MA.

Raymond Choo, K.-K. (2008). Politically exposed persons (PEPs): Risks and mitigation. *Journal of Money Laundering Control*, 11(4):371–387.

Rijkers, B., Baghdadi, L., and Raballand, G. (2015). Political connections and tariff evasion evidence from Tunisia. *The World Bank Economic Review*, 31(2):459–482.

Robinson, J. A., Torvik, R., and Verdier, T. (2006). Political foundations of the resource curse. *Journal of Development Economics*, 79(2):447–468.

Svolik, M. W. (2012). *The politics of authoritarian rule*. Cambridge University Press, Cambridge.

Szeidl, A. and Szucs, F. (2017). Media capture through favor exchange. Discussion Paper DP11875, CEPR.

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

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440.

World Bank (2019). *World Development Indicators*. The World Bank, Washington, DC.

A Additional Tables

Table A1: Country-level regressions of largest component share using lag democracy and income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std(Economist$_{2007}$)	0.077***				0.079***			
	(0.014)				(0.017)			
Std(Freedom House$_{2007}$)		0.055***				0.051***		
		(0.014)				(0.016)		
Std(Polity IV$_{2007}$)			0.041***				0.034**	
			(0.013)				(0.014)	
Std(V-DEM$_{2007}$)				0.059***				0.054***
				(0.014)				(0.016)
Std(log(GDP per capita$_{2007}$))	0.023	0.040***	0.055***	0.034**	0.021	0.033**	0.045***	0.027*
	(0.015)	(0.015)	(0.013)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)
Std(log(Rents))					0.010	-0.003	-0.009	-0.004
					(0.019)	(0.018)	(0.019)	(0.018)
Std(log(Nodes))					0.015	0.020	0.019	0.019
					(0.016)	(0.016)	(0.017)	(0.016)
Dependent Mean	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457
Observations	153	153	153	153	153	153	153	153
R^2	0.260	0.203	0.175	0.206	0.268	0.214	0.187	0.217
Adjusted R^2	0.250	0.193	0.164	0.196	0.248	0.193	0.165	0.195

Notes: All independent variables are standardised, so the coefficients represent the change in the dependent for one standard deviation increase in the respective variable. Standard errors in the parentheses are robust to heteroskedasticity and serial correlation. Significance levels *p<0.1; **p<0.05; ***p<0.01