Early Determination of Career Advancement?
Evidence from Patronage Networks in the Chinese Bureaucracy

Jason Luo[1]

November 30, 2019

Abstract

What determines political promotion in authoritarian regimes? Do autocrats sort on the competence or political loyalty of their subordinates? A long-running scholarly debate in both economics and political science concerns whether the selection of leaders in the Chinese bureaucracy is meritocratic or personalistic. Using the latest demographic, career experience, and economic performance data for more than four thousand leaders starting from the prefecture-level in China from 1995-2015, I find that the career outcomes of political leaders are largely determined by patronage networks formed in early periods of their career. Early patronage relationships remain the best predictor of a leader's later career advancement even after controlling for economic performance. Adopting a novel network-based measurement strategy, this study goes beyond existing one-to-one, binary coding schemes of political connection to capture neighborhood views of leaders in the network. To the best of my knowledge, this paper is one of the first to use advanced machine learning and network analysis methods to quantify the network impact of political patronage on promotion and to compare it with performance among nationwide politicians in large bureaucracies.

[1] Department of Political Science, Stanford University. Contact: jasonluo@stanford.edu.

I. Introduction

Why do some politicians rise to the top while others do not? Is the selection of leaders in non-democratic countries determined by their performance, competence, and quality? Or are they determined, instead, by their connection, loyalty, and positioning in political factions and patronage networks?

In this study, I address the above questions by examining the linkage between patronage networks, economic performance, and political promotion in contemporary China. The central task of the paper is to map out patronage networks and compare the predictive power between early-stage patronage networks and economic performance on leaders' advancement in later career periods. Based on extensive demographic, career, and economic performance information for more than four thousand contemporary Chinese leaders at the prefectural-level and above from 1995 to 2015, this paper finds that a network-based deep learning model is able to predict career outcomes at the end of the study period with a high accuracy rate. Even after controlling for leaders' ability and performance in generating economic growth, patronage relationships formed in early periods – specifically until ten years prior to 2015 – remain the best predictor of their promotion outcome in 2015. This finding suggests that political patronage outweighs economic performance in career advancement of middle-level politicians in contemporary China. In this paper, I also develop a network-based measurement strategy of political patronage by adopting a cutting-edge feature learning algorithm in computer science, which goes beyond existing one-to-one, binary coding scheme of political connection to capture neighborhood views of leaders in patronage networks and allows for high-quality predictions for promotion.

Section II reviews the existing literature and discusses the contributions of this study. In Section III, I outline the cadre evaluation and promotion system in China as background and raise two competing hypotheses derived from existing theories. Section IV provides a detailed description of the patronage and performance data and introduces network analysis methods to build three types of network graphs as the basis for subsequent analysis. Section V presents measurement strategies of patronage relationships and economic performance from existing data. Section VI estimates four classes of model to predict promotion and compares predictive power between patronage and performance. Section VII concludes by discussing the limits of this study and by suggesting future avenues of research.

II. Literature Review

This paper connects most directly to the following three bodies of literature. First and foremost, it speaks to a key debate in economics and political science about the determinants of political promotion and associated economic outcomes in authoritarian regimes like China. Ever since Max Weber ([1921] 1958) made the distinction between rational, impersonalized management and patriarchical management filled with loyal partisans, political scientists have focused on analyzing the meritocratic and clientelist types of political systems (e.g. Oi 1985 & 1999; Frye and Shleifer 1997; Stokes 2005). Particularly, a long-standing, large group of research has discussed leadership selection in Chinese politics. Early works on promotion in the Chinese bureaucracy are concerned mostly with procedures, metrics, and exit patterns of career mobility and their institutional implications for governance of the communist bureaucratic system (Oksenberg 1976; Manion 1985; Zhou 2001; Whiting 2004). An important work of Walder (1995) proposes a model of two distinct elite selection paths along institutionalized – administrative and professional – career lines that lead to a divided elite. Zhao and Zhou (2004) verify and enrich Walder's dual-path model for promotion and find that the value of educational credentials has increased in the Reform Era. More recently, Zhou (2016) and Zhou et al. (2018) have theorized and empirically tested the "stratified mobility" patterns in the Chinese bureaucracy and argued that such personnel management practices have provided stable institutional bases for central-local government relationships and set limits to the downward reach of the state and the upward reach of local interests.

Scholars of Chinese politics have fiercely debated whether the primary determinant of leadership promotion is performance or patronage (e.g. Chen and Kung 2016; Chen, Li, and Zhou 2005; Li and Zhou 2005; Opper, Nee, and Brehm 2015; Shih, Adolph, and Liu 2012; Yao and Zhang 2015; Keller 2016; Landry 2008; Lü and Landry 2014; Tao et al. 2010). Scholars advocating for the performance-based view have explained China's remarkable growth in the past three or four decades as a result of merit-based competition among elites. These include the seminal contributions of Qian and Weingast (1997), Maskin, Qian, and Xu (2000), Chen, Li, and Zhou (2005), Jin, Qian, and Weingast (2005), and Guo (2007), in which they argue that an extensive, impartial, performance-based cadre evaluation system, combined with a geography-based, multidivisional form bureaucracy, provides strong incentives for regional administrators to compete with each other to generate high economic growth in order to advance their career in the Chinese bureaucratic system. Following a similar logic of inter-jurisdictional competition, local officials are also incentivized to submit more taxes to their superiors in order to signal their performance and to stand out in the game of promotion (Lü and Landry 2014). Also, the latest work of Landry, Lü, and Duan (2018) shows that performance contributes to the promotion of county-level cadres, but not for prefecture and provincial leaders, suggesting that economic performance may play a greater role in promotion at lower administrative levels of government than at higher ones.

In contrast to the performance-based view of leader selection, proponents of the patronage argument emphasize political connections as the key driver, for autocratic rulers likely place a greater priority on maintaining short-term state capacity and buying the support of winning coalitions than on providing broadly encompassing goods such as economic growth (Bueno de Mesquita et al. 2003; Gandhi and Przeworski 2006). Based on exhaustive biographical data on CCP's Central Committee members, Shih, Adolph, and Liu (2012) and Meyer, Shih, and Lee

(2016) have found no evidence that strong growth performance was rewarded with higher party ranks at any of the post-reform party congresses. Instead, factional ties with various top leaders played substantial roles in elite ranking. Moreover, quantitative studies of lower-level leaders do not find much evidence for growth-based promotions. At the city level, for example, Landry (2008) finds that exceptional economic performance has almost no effect on the likelihood that any given mayor is internally promoted to the position of party secretary. Although Lu (2018) finds patronage sponsorship to have positive effects on leader promotion at all levels of government within Jiangsu Province, the author does not measure or control for economic performance.

Recent studies have also tried to move beyond this dichotomic view of patronage and performance to show the possible coexistence and interaction between the two. In his important work, Junyan Jiang (2018) advances a view that patronage networks promote the economic performance of local governments in China, where he shows that city leaders with informal ties to the incumbent provincial leaders deliver significantly faster economic growth than those without such ties. This implies patronage could also have a more subtle way of influencing political selection that works through, rather than against, the formally meritocratic institutions. In another important work, Jia, Kudamatsu, and Seim (2015) argue that patronage ties may facilitate meritocratic selection by giving senior leaders greater confidence in the loyalty of their subordinates, and they show that there exists a complementary relationship between political connection and performance in predicting provincial-level promotions.

Despite a huge literature on political promotions in China, direct and systematic examination of the relative importance of patronage and performance in driving elite selection is scarce and the current one-to-one, binary measurement strategy of political connection does not allow for the use of abundant information encoded in patronage networks. We still do not have a clear picture of the relative importance of patronage and performance in the selection of leaders, perhaps due to both the limited availability of a nationwide database and the lack of a good measurement strategy for analyzing patronage networks. The following table summarizes the questions, findings, data, and patronage measurement strategies of recent quantitative works on promotion in the Chinese bureaucracy. I will discussion the content of the last column, "Patronage Measurement," later in this section when I explain how my paper departs from the existing literature.

Table 2.1: Summary of Quantitative Literature on Promotion in Chinese Politics

Study	Question or Focus	Data	Finding	Patronage Measurement
Zhou (2001) & Zhao and Zhou (2004)	The recruitment and promotion patterns of socialist bureaucrats	A representative sample of 5,000 urban residents drawn from a multistage scheme in 20 cities in China in 1993 and 1994. Note this is not on leaders, but on ordinary people.	Varying selection criteria over time and two distinct patterns of promotion between national bureaucratic systems and within workplaces.	None (not applicable)
Li and Zhou (2005)	Driver of promotion for provincial leaders	254 provincial leaders who served in 28 Chinese provincial units from 1979-1995	Positive correlation between performance and promotion	Not really taken into account. A close variable is whether a provincial leader has previous experience or holds joint-appointment in the central government

Guo (2007)	Driver of promotion for county leaders	Nationwide county leaders from 1995-2002	Positive correlation between promotion and government revenue growth, but not economic growth.	None
Landry (2008)	Internal promotion of mayors to the position of prefecture-level party secretary	Mayors of 104 cities from 1990-2001	No correlation between performance and promotion	None
Shih, Adolph, and Liu (2012) & Meyer, Shih, and Lee (2016)	Driver of rankings of the Central Committee members	Top leaders in the 12th through 16th Party Congresses (1982–2007).	No correlation between performance and ranking. Positive correlation between factional ties and ranking.	One-to-one factional tie between a regular leader and a top leader. Based on shared birthplaces, educational institutions or work units.
Lü and Landry (2014)	Effect of political competition on fiscal revenues (taxation)	Nationwide counties from 1999-2006. Note the unit of study is not a leader, but a county.	Strong inverse U-shaped relationship between political competition and taxation	None (not needed)
Zhou (2016) & Zhou et al. (2018)	"Stratified mobility" – the bureaucratic personnel flow pattern among officials across levels of administrative jurisdiction.	Jiangsu Bureaucratic Structure data (JSBS) that contains about 41,000 officials' career information from all levels of government within Jiangsu province from 1990-2008.	Some key empirical patterns of bureaucratic personnel flow and mobility among offices and across administrative levels.	None but has intensive discussions about patterns and properties of patronage networks based on overlapping work experience. Note that these two studies do not look at impact of performance or patronage on promotion.
Lu (2018)	The effect of patronage on promotion for all county-level, prefecture-level,	Jiangsu Bureaucratic Structure data (JSBS) as above.	Positive effect of political connection/sponsorship on promotion, while	One-to-one connection between a client to an incumbent/previous leader based on

	and provincial-level leaders within Jiangsu Province		that effect can turn negative after a patron loses power.	shared work experiences. Note that there is no economic performance measure for officials in this study, perhaps because most of those in her sample are not a head of some level of government and attribution is hard.
Landry et al. (2018)	Driver of promotion for provincial, prefectural, and county leaders	Nationwide provincial, prefectural, and county-level leaders from 1999-2007	Mixed results. Positive correlation between performance and promotion at county level, but not prefecture-level and above	One-to-one appointment-based connection where a client was promoted within the same work unit when he/she worked under the patron
Jiang (2018)	The effect of patronage on economic growth	Nationwide city-level panel data from 2000-2011; nationwide prefecture leaders and above from 1995-2015	City leaders with informal ties to the incumbent provincial leaders deliver significantly faster economic growth than those without.	One-to-one appointment-based connection where a client was promoted within the same work unit when he/she worked under the patron
Keller (2016)	Proposing network analysis approach to analyze Chinese leaders	Central Committee members from 1982-2002	Positive effects of patronage networks on promotion. No control for economic performance.	One-to-one connection based on shared hometown, career, and school experiences. Also uses network centrality measures
This paper (Luo 2019)	Driver of promotion for leaders at prefecture-level and above	Nationwide prefecture-level leaders and above from 1995-2015	Strong positive effects of patronage networks on promotion, even after accounting for performance	Identification of connection is similar: based on shared hometown, career, and school experiences; But uses a new measurement strategy that moves

				beyond the existing one-to-one, binary coding to capture neighborhood views of leaders in patronage networks.

Secondly, this paper also complements a growing body of research concerning the quality of political selection around the world and the potential trade-offs it entails. While autocrats face a trade-off between competence and political loyalty when promoting subordinates (e.g. Egorov and Sonin 2011), representative democracy can also struggle to deliver both high-ability politicians and broad representation (e.g. Caselli and Morelli 2004; Dal Bó et al. 2017). The novel approach proposed by this paper to study the dynamics of promotion can thus add to these studies and provide future researchers a broadly applicable network-based method to compare political selection across autocracies and democracies around the world.

Thirdly, this study joins the long tradition of studying patronage relationships and clientelism in authoritarian regimes (Scott 1972; Oi 1985). While previous studies mostly probe into one-to-one relationships between patrons and their clients, or many-to-one connections for members belonging to political factions (Li 2001; Miller 2004), a growing body of research advances a social network approach in examining patronage relationships (Razo 2008; Keller 2016; Lu 2018). By adding one-to-many and many-to-many perspectives to study patronage, this network-based approach goes beyond the simple aggregation observation that patrons promote clients and allows for rigorous, quantitative exploration of "*an informal institution that reproduces itself and a network in which patrons are connected to clients, who in turn act as patrons to other clients* *on a lower level*" (Keller 2016). Besides solving the patronage identification problem and examining career tracks, a growing body of research has also pointed to the consequences of patronage other than political selection per se. Jiang and Zeng (forthcoming), for example, demonstrate that city-level reforms associated with land-related mass grievances only occur among a subset of city leaders with informal connections to the higher-level authority. This emphasizes the importance of access to external patronage and support networks for improving local politicians' responsiveness to the ordinary citizen.

This paper departs from previous studies in four important ways. First, rather than viewing patronage as a binary connection to either a specific leader (one-to-one), or to a particular political faction (still one-to-one for the client), I treat patronage relationships from a network-based perspective (one-to-many and many-to-many). Specifically, for every politician in my sample, I use home origins, overlapping work and school experiences, and subnational personnel appointment information to map out the existence, direction, and strength of patron-client links between this leader and all others in the database. This leads to the construction of a SNAP[1] network graph of 3,844 nodes and 523,027 directed, weighted edges, with each node being one official and edge being patron-client link.

Second and most importantly, this paper uses a new measurement strategy of patronage that moves beyond the existing one-to-one, binary coding of political connections to capture neighborhood views of leaders in patronage networks. Scholars have tried various ways to

[1] Stanford Network Analysis Platform (SNAP) is a general-purpose network analysis and graph mining library. It efficiently manipulates large graphs, calculates structural properties, generates regular and random graphs, and supports attributes on nodes and edges. It is widely used and among the best network analysis tools.

identify political connections, including faction-based grouping (e.g. Shih, Adolph, Liu 2012), overlapping hometown, school, and career experiences (e.g. Jia et. al 2015), and appointment-based connection where a client was promoted within the same work unit when he/she worked under the patron (e.g. Keller 2016, Jiang 2018, and Landry et al. 2018). However, regardless of their differences in identification strategy, all current research, to the best of my knowledge, in essence treat and code patronage as a binary variable: that is, a client is either connected to a patron/faction or not. The only exception is Keller (2016), where the author uses a number of network centrality metrics, but those are too abstract to capture the complexity of neighborhood connectivity among leaders. The new measurement strategy adopted in this paper allows for numerical representation of both local neighborhoods (micro-views) and global networks (macro-views) of each leader in the entire nationwide patronage graph. In other words, while still relying on similar identification of connection, this new measurement strategy captures what local patronage networks for a leader look like, with a flexible notion of defining neighborhoods. In doing so, this paper extracts complex information encoded in patronage networks and uses it to predict promotions. This new measurement is done through a cutting-edge feature learning algorithm in computer science, "*Random Walks (node2vec),*" which captures the diversity of connectivity patterns observed in patronage networks in the Chinese bureaucracy.

Third, this study also improves from a substantial part of existing research based on insider sources of political connection by drawing from the latest systematic, nationwide, and publicly available biographical information through official channels of the Chinese government. Lastly, instead of focusing on factions, struggles, and conflicts among national leaders, i.e. Politburo and Central Committee members, I mostly look at the career pathways of municipality administrators and their early interactions with colleagues, leaders, and subordinates. Such a focus down the bureaucratic ladder includes a larger pool of middle-level political elites for a more systematic test of the relative importance of patronage and promotion in leader selection in China.

In summary, the paper's most important departure from existing literature is the use of a new network-based measurement strategy that allows for the comprehensive representation and quantification of a leader's patronage neighborhood views for predictive inference. To the best of my knowledge, this study offers the first network-based analysis of nationwide middle-level leaders and above in contemporary China to test the relative importance of performance and patronage in driving promotion.

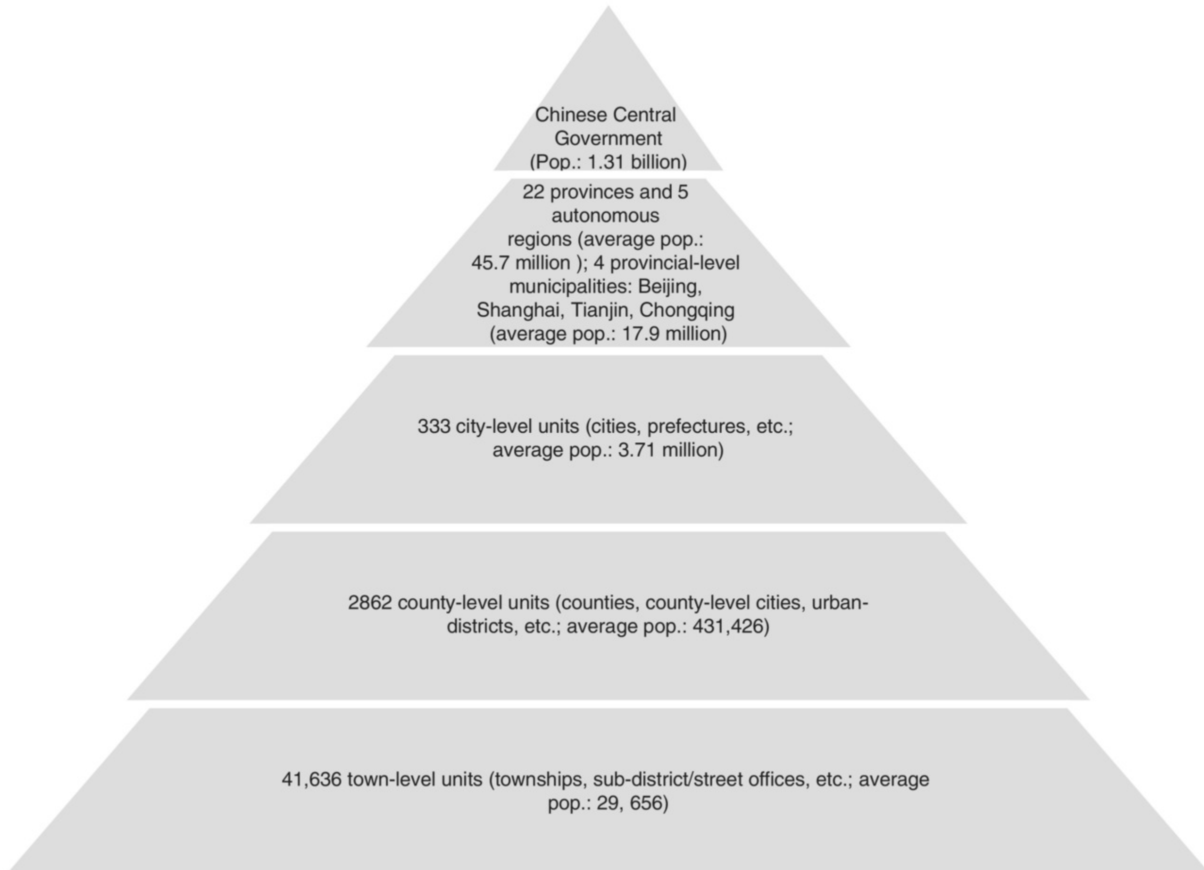
III. Background and Hypotheses

i. Cadre Management and Promotion in the Chinese Bureaucracy

Similar to their counterparts in the former Soviet Union and other Eastern Bloc countries, the Communist Party of China (CPC) embeds itself fully into the Chinese state and adopts a *nomenklatura* cadre management system in which the selection of political leaders is solely determined by leaders one level immediate above them (Burn 1987). China has four levels of governments: the central/national government, the provincial governments, the prefecture governments, and the county governments. Therefore, the *nomenklatura* system works in a top-down fashion whereby leaders at the national-level decide provincial heads, provincial-level leaders appoint prefecture-level (cities and municipalities) top cadres, and prefecture heads determine the county leadership. The People's Republic of China rules thirty-three province-equivalent jurisdictions, including provinces, direct-controlled municipalities, and ethnic minority autonomous regions. A medium-size Chinese province consists of about ten to twelve prefecture-level cities and hosts around 45 million people. The following figure shows a vertical view of levels in the Chinese government and associated population sizes.

Figure 3.1 The Structure of the Chinese Government[1]

[1] Source: Xu Chenggang, The Fundamental Institutions of China's Reform and Development, Journal of Economic Literature 2011, 49:4, 1076–1151, 1084.

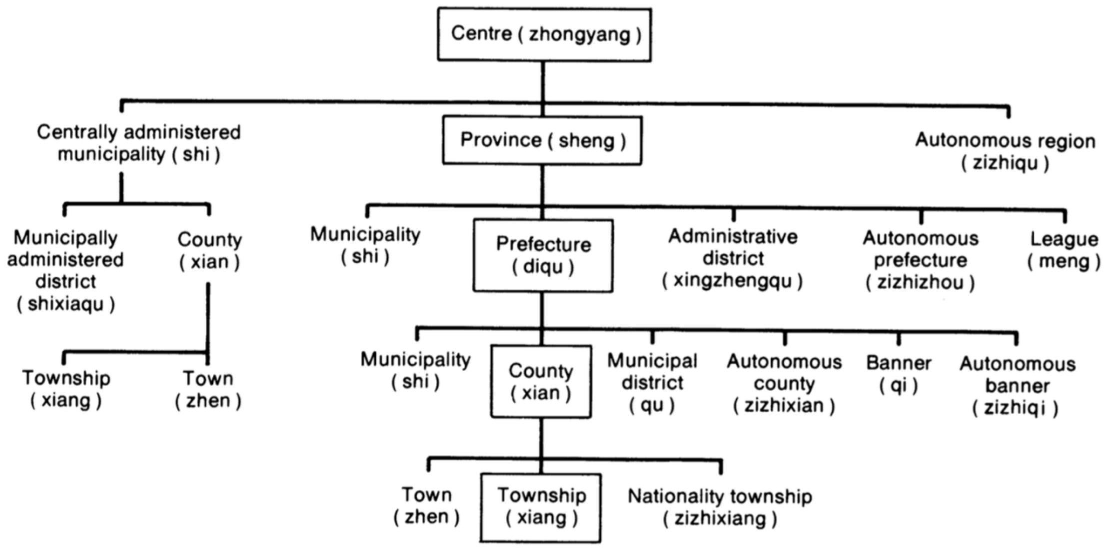


Throughout China, party and government structures closely parallel one another, with party committees and representatives present in local government agencies. Party secretaries are more powerful than the heads of the government at the same level (governors, mayors, and magistrates) and usually have the final say on important governmental decisions (Jia and Xu 2018). According to the 1983 official handbook, the fundamental principle of cadre management is, and always has been, Party management.[1] In principle, no leader's appointment, promotion, transfer or removal can be effective without the approval of the appropriate Party committee (Manion 1985). For the management of leading cadres, the most important Party committees are

[1] Source: Central Organization Department, 1983 Handbook, pp. 88. Cited from Manion (1985).

those which are territorially based – the Central Committee nationally and the local Party committees down to the county level. The following figure shows the territorial levels.

Figure 3.2 The Territorial Levels for Cadre Management[1]



The domain of this study is leaders who have made it to the level of Prefecture heads and above, which corresponds to the top three rows in the above figure: the Center, the Province, and the Prefecture. The choice results from the availability of high-quality official data but is also made to focus on most important and relevant players in the game of political promotion in China. The following table gives a list of official ranks in the Chinese party-state. Again, due to the nature of the database, which will be explained in the Data Section, the leaders included in the sample of this study are those from Level 5 – Mayors and Party Secretaries of prefectural party committees – and above.

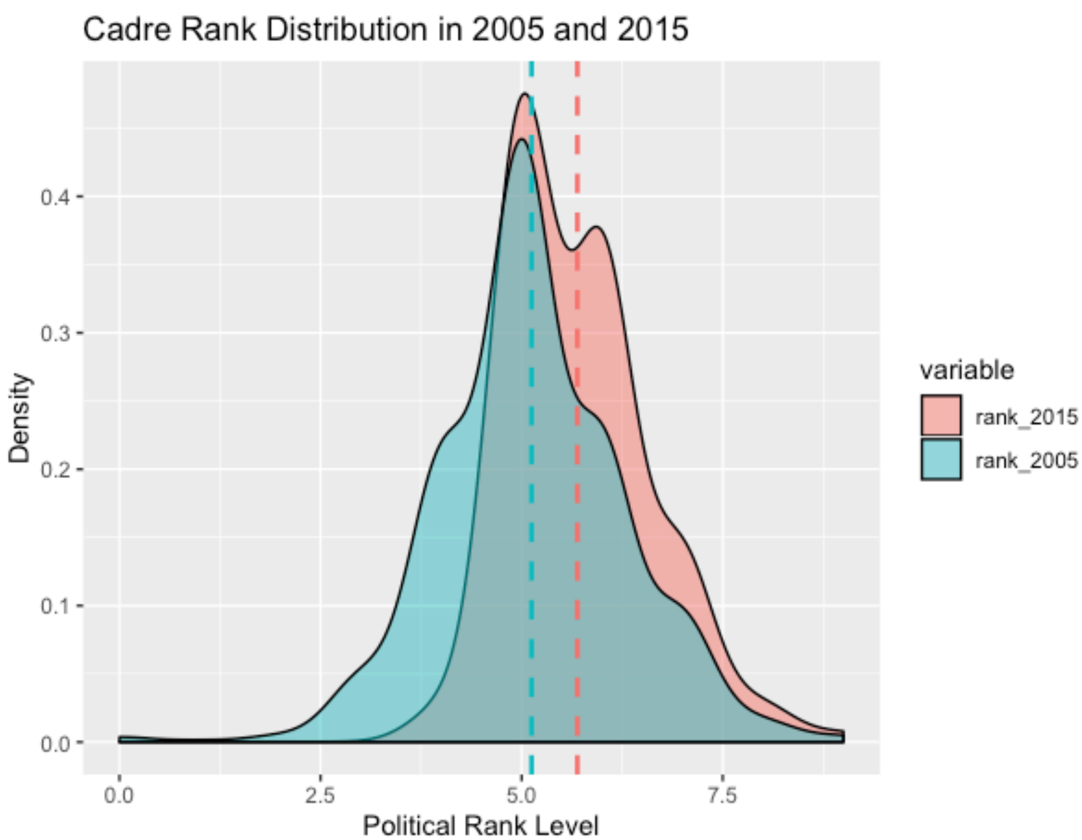
[1] Source: Manion, Melanie. "The cadre management system, post-Mao: The appointment, promotion, transfer and removal of party and state leaders." The China Quarterly 102 (1985): 203-233.

Table 3.1: Official Cadre Ranks and Positions in China

Cadre Rank Level	Party Positions	Government Positions
9: Central/National Leader 正国级	Secretary-General of Central Committee, Politburo Standing Committee Members	Premier of State Council, Secretary-General of National People's Congress, etc.
8: Deputy Central/National Leader 副国级	Politburo members, Secretary of General Office, etc.	Vice Premiers, Head of the Supreme Court, Procurator-General
7: Provincial/Ministerial Head 正部	Deputy secretaries, Heads of departments and commissions, Secretaries of provincial committees	Heads of ministries, Governors of provincial governments
6: Deputy Provincial/Ministerial Head 副部	Deputy secretaries of provincial committees, secretary of discipline inspect commission and other provincial standing committee members	Vice governors of provincial governments, Head of the people's Court, Chief Procurator
5: Bureau Director / Prefectural Head 正厅	Secretaries of perfectural committees, Heads of provincial departments and commissions, other members of provincial committees	Mayors (Heads) of prefecture/city/municipality governments, Heads of provincial-level bureaus, chairs of prefectural people's congresses
4: Deputy Bureau Director / Prefectural Deputy Head 副厅	Deputy heads of provincial departments and commissions, members of prefectural standing committees	Vice mayors of prefecture/city/municipality governments, Vice heads of provincial-level bureaus
3: Division/County Head 正处	Heads of prefectural departments and commissions, other members of prefectural committees	Heads of county and city district governments, heads of prefectural government divisions, chairs of country people's congresses
2: Deputy Division Head 副处	Deputy heads of prefectural departments and commissions, members of county standing committees	Vice heads of county and city district governments, Vice heads of prefectural government divisions.
1: Below Deputy Division Head 小于副处	County party committee members, etc.	Township heads, etc.

The empirical design of this paper is intended to examine changes in cadre rank from 2005 to 2015. The outcome variable of the paper, *final rank*, is a leader's cadre level in the Chinese bureaucracy in 2015. If a leader retired, died, or was prosecuted before 2015, I record his/her highest rank before exiting. Below is a density plot comparing rank distributions in 2015 and in 2005. It is clear that after ten years, the entire rank distribution shifts to the right, with the mean changing from 5.1 (in 2005) to 5.7 (in 2015), thus suggesting a general increase in ranks among the leaders in the dataset. My goal is to compare the relative importance of patronage and performance information in explaining this part of the variation, i.e. the changes in cadre rank for all leaders from 2005-2015.

Figure 6.2: Cadre Rank Distributions in 2005 and in 2015



ii. Existing Theories of Promotion in China

The Performance Camp: The post-Mao leadership formalized a comprehensive evaluation system of local cadres, which served to keep track of their administrative performance on a series of major "social, economic, and cultural" targets.[1] Whiting (2004) emphasizes the important transition of moving away from what was seen as subjective evaluations of political attitudes toward specific, measurable, and quantifiable indicators of performance. At the county level, these measures ranged from the gross value of industrial output to tax remittances and procurement of agricultural and agricultural subsidiary products, and from realized investment in infrastructure to the population growth rate and the completion rate for nine-year compulsory education.[2] Furthermore, there are also political "priority" targets with veto power (yipiao foujue), which means if leaders fail to attain these targets, this would cancel out all other work performance metrics, however successful, in their comprehensive evaluation at the end of the year (Edin 2003). The existence of a comprehensive system for scoring officials on their policy performance seems to suggest that the Chinese elite was motivated by the scoring system to perform well (Edin 2003; Landry 2008; Whiting 2004).

The Performance camp views that among all meritocratic criteria, the most important driver of career advancement is economic performance. Proponents of the Performance camp have argued that an extensive, impartial, performance-based cadre evaluation system, combined with a

---

[1] Whiting (2004) offers a most thorough account of the evaluation of the cadre management system since the reform period and the details of official criteria and de facto factors.

[2] Source: Zhonggong zhongyang zuzhibu, "Guanyu shixing difang dangzheng lingdao ganbu niandu gongzuo kaohe zhidu de tongzhi" ("Notice Regarding Implementation of the Annual Job Evaluation System for Leading Cadres of Local Party and Government Organs"), Zhong- guo renshi nianjian (Beijing: Zhongguo renshi chubanshe, 1991). Cited from Whiting (2004).

geography-based, multidivisional bureaucracy, provides strong incentives for regional administrators to compete with each other to generate high economic growth in order to advance their career in the Chinese bureaucratic system (e.g. Chen, Li, and Zhou (2005), Li and Zhou (2005), Maskin, Qian, and Xu (2000), and Jin, Qian, and Wingast (2005)). This is a game Li (2007) coins "*Promotion Tournament*" and is meant to explain the remarkable growth of the Chinese economy in the past three or four decades.

The Patronage Camp: In contrast to the performance view, scholars have argued that autocrats face a trade-off between competence and political loyalty when promoting subordinates (Egorov and Sonin 2011), and they likely place a greater priority on maintaining short-term state capacity and buying the support of winning coalitions than on providing broadly encompassing goods such as economic growth (Bueno de Mesquita et al. 2003; Gandhi and Przeworski 2006). An extensive literature on China has documented that political connections are conducive to career advancement, including seminal contributions such as Guo (2007), Landry (2008), Shih, Adolf, and Liu (2012), and Jia, Kudamatsu, and Seim (2015).

A key argument of the Patronage camp is that political connections, being the dominant driver of promotion, can work through economic performance in a way that better connected leaders are placed *ex ante* by their patrons on positions that are more likely to yield better GDP growth. In other words, the strong correlation between performance and promotion, if observed, may be attributed to the selection bias resulted from omitting the underlying political connections variable at work.

From these two camps, I derive the following two competing hypotheses:

> *Hypothesis I (the Performance camp):* Economic Performance is the dominant criterion driving elite advancement. The ability of leaders to generate GDP growth remains the best predictor of promotion, even after measures of patronage are controlled for.

> *Hypothesis II (the Patronage camp):* Connection to patronage networks and strength of patron-client relationships is the *de facto* driver of political selection. In other words, once it is taken into account, the predictive power of economic performance goes away for career advancement.

What follows is the introduction of my data, methods and models, which will allow me to test which of the above hypotheses is closer to the truth.

IV. Data and Methods

i. Patronage Data and Network Graphs

The main source of patronage data in this paper is from the Chinese Political Elite Database (CPED), a large biographical database that contains extensive demographic and career information of over 4,000 key city, provincial and national leaders in China since the mid 1990s. It is constructed and maintained by Junyan Jiang. The database intends to provide standardized, accessible information for empirical researchers who are interested in Chinese leadership and is supported by grants from the Social Sciences Division at the University of Chicago, the Ford Foundation, and the National Science Foundation (SES‐1560513). It was made available along with the paper, *"Making Bureaucracy Work,"* published in 2018 on AJPS. Since the database codes objective demographic and career experience information from official publications, the existence of non-random errors that lead to systematic biases is highly unlikely.

Currently, the database includes:

- All city party secretaries, mayors (including party and government heads under centrally administered municipalities) between 2000 to 2015
- All civilian members of provincial standing committee from 2000 to 2012
- All provincial governors and provincial secretaries from 1995 to 2015
- Full and alternate members of 15th to 18th Central Committee (1997-2012)

For each leader, the database provides information about the time, place, organization, and rank of every job assignment listed in their curriculum vitae, which is collected from government websites, yearbooks, and other authoritative online sources. The final output from the database contains two separate tables on the official's time-invariant attributes and time-varying career information.

Figure 4.1: Heat Map of Hometown Origin of All Politicians in the Dataset. Xinjiang has 78 leaders.

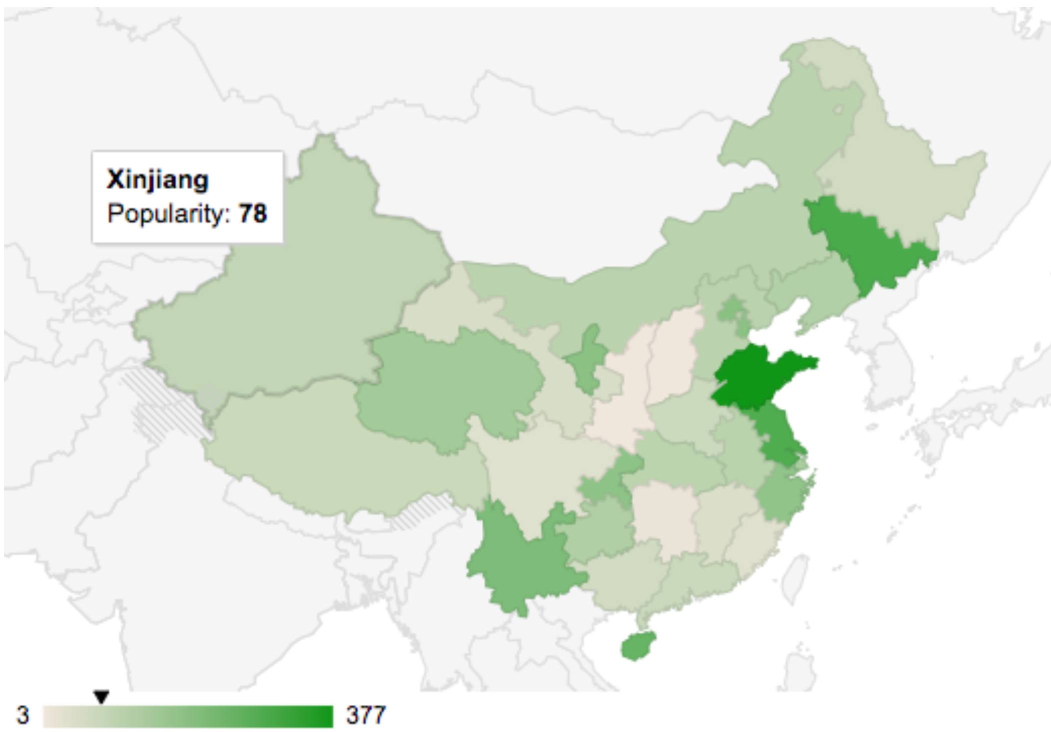


Figure 1 shows the distribution of politicians in the dataset according to their home origins. Most leaders are from the east coast, whereas a handful of politicians come from the middle of China. This is largely in line with the population distribution in China. The following Table 1 gives more details about the database, including the counts of politician, province, city, and also the total number of entries in the work experience sheet.

Table 4.1: Basics of the Database

Item	Count
Leaders	4,057
Provinces	32
Prefectures	389
Work Experience Data Points	62,742

The next table presents summary statistics of demographic attributes of leaders in the database.

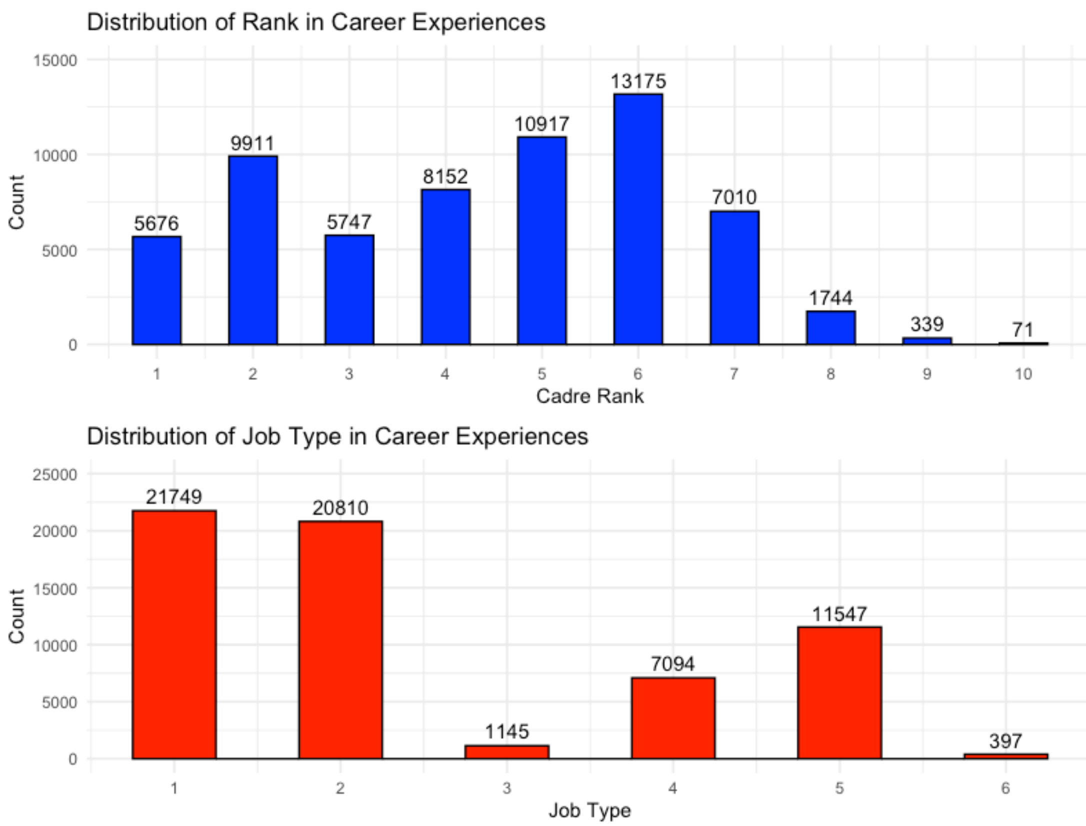
Table 4.2: Summary Statistics for Political Leaders

Covariates	Mean	SD	25th Percentile	75th Percentile
Gender (Male)	0.94	--	--	--
Ethnicity (Han)	0.89	--	--	--
Education Level	4.67	0.86	4	5
Cadre Rank	5.68	0.89	5	6
Currently In Office	0.59	--	--	--
Age (in year)	60.72	8.55	53.78	65.96
Party Membership Years	36.97	9.19	30.35	41.92

Note:
1. All information is updated up until July 1, 2015.
2. N =3,918. Leaders with missing information were removed from analysis.
3. Codes for Education Level: 0: NA 不详, 1: Middle School 初中, 2: High School 高中, 3: Associated Degree 专科, 4: Bachelor's Degree 本科, 5: Master's 硕士, 6: PhD 博士.
4. Codes for Cadre Rank: 0: No Level, 1: Below Deputy Division Head 小于副处, 2: Deputy Division Head 副处, 3: Division Head 正处, 4: Deputy Bureau Director 副厅, 5: Bureau Director 正厅, 6: Deputy Ministerial Head 副部, 7: Ministerial Head 正部, 8: Deputy National Leader 副国, 9: National Leader 正国.
5. Out of the 120 *military* leaders in the database, nearly all of them miss career information (N=119). The analysis in this paper thus focuses on *civilian* politicians and includes only one PLA leader.

In addition to attributes at the leader-level, the following graph summarizes the distributions of cadre ranks and job types at career experience level.

Figure 4.2: Rank and Job Type Distributions in Career Experience



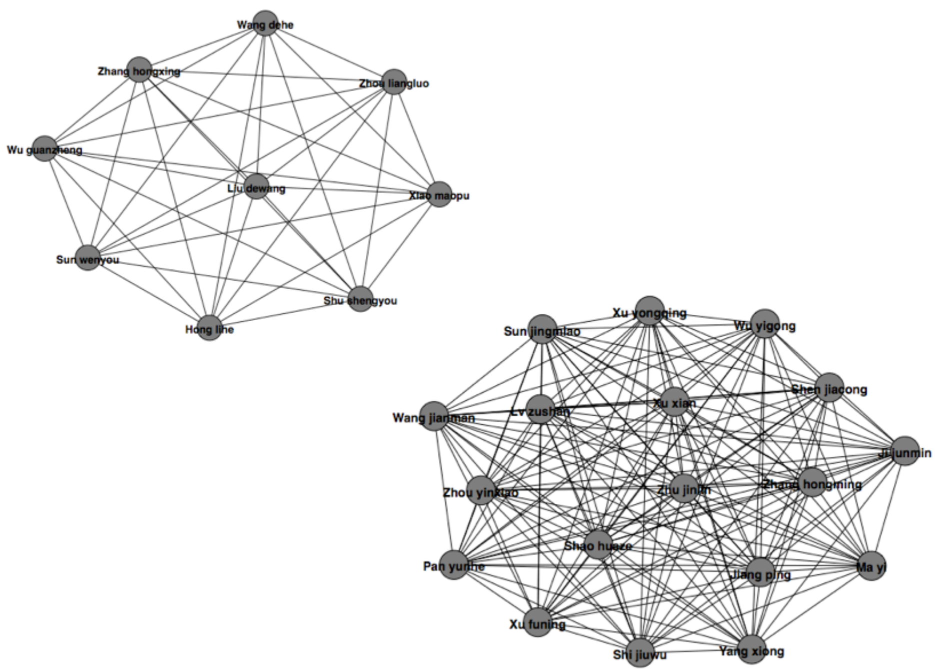
Note: Codes for Job Type are 1: Party Committee/Youth League 党委/共青团, 2: Government 政府, 3: Military 军队, 4: People's Congress/Political Consultative Conference/Business 人大/政协/企业, 5: School 学校, 6: Others 其他.

The second part of this section presents how I use the plain text database to build social network analysis (SNA) graphs. I construct three main network graphs to approximate the patronage network in real life, including hometown origin network, work/school experience overlap network, and promotion network.

A. Home Origin Network

Obviously, the hometown network is highly clustered around every city. For each city cluster, all nodes are fully connected by construction. There are 59,306 edges in the network. Two nodes are connected if the leaders they represent are from the same city. Each edge thus represents a pair of politicians coming from the same city. The following graph presents the subnetworks for Hangzhou (bottom right) and Shangrao (top left).

Figure 4.3: Leader Clusters for Hangzhou and Shangrao in Home Origin Network

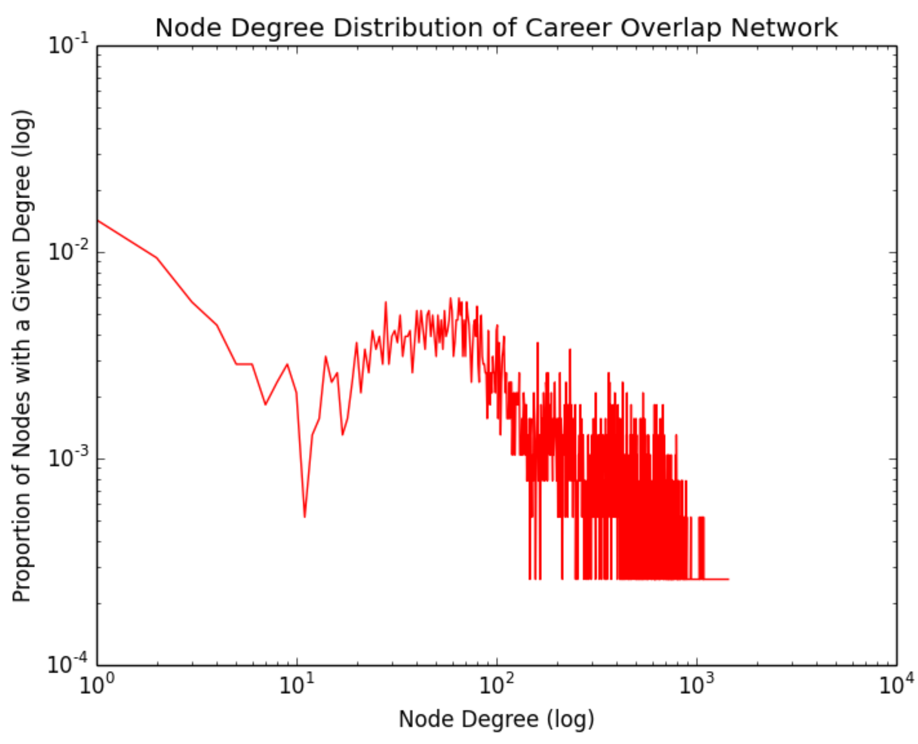


B. Career Experience Overlap Network

I next construct a directed, weighted graph based on overlapping work and school experience between the 4,057 political leaders. After dropping politicians with missing career information, this results in a network graph of 3,844 nodes and 523,027 directed, weighted edges.

An edge is firstly formed if two leaders have worked together in the same municipality (of course, in the same province) under the same job type for a time period longer than six months. I consider it to be one overlap. This is a rather rigorous measure of overlapping career experience, for two leaders have to not only work in the same prefecture for over 180 days, but also within the same branch of the Chinese party-state, e.g. party, or government, or school, etc. Multiple overlaps, if they exist, are then factored into the edge weight, which indicates the total years of overlapping work experience. Lastly, the direction (and the strength of the direction) of an edge is recorded through a comparison of cadre ranks of the two leaders averaged over all the time periods of their co-working experience.

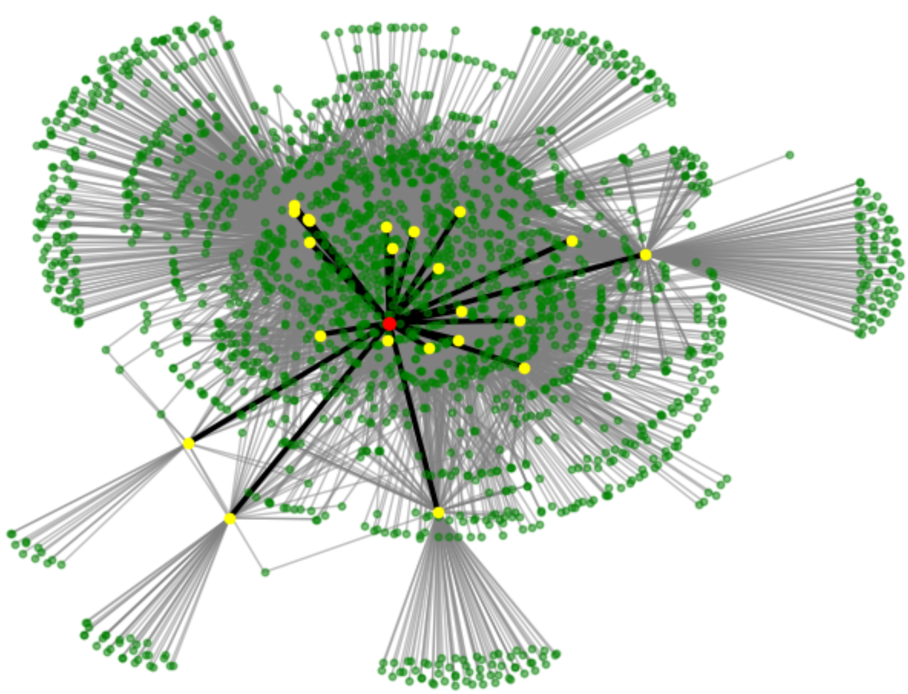
Figure 4.4: Degree Distribution of Nodes in the Overlap Network



The above graph gives the node degree distribution of the overlap network, which suggests that most nodes have degrees between 100 and 1000. That is, the majority of leaders in our overlap network have worked together with one hundred to one thousand politicians in the network. This indicates that leaders are widely connected among themselves, although the strength of connections varies (i.e. strong and weak ties).

I now turn to a specific leader. I randomly pick one leader, node #3057, named Zhao Zhuping. He is currently the government head of Minhang District of Shanghai, a cadre rank equivalent to a municipality head or a county top leader in United States. Interestingly, he was also trained at the Harvard Kennedy School for a short period in 2009 for public administration related issues. He has only 76 patron-client links in the network, which is much lower than the average node degree (272.1), partly because he is relatively young (51 years old) and also because he serves a moderate position in the bureaucratic system. The following graph shows the top neighbors of node #3057 a.k.a. Zhao Zhuping and the neighbors of these top neighbors.

Figure 4.5: Two-Hop Neighbors of Node 3057 (Zhao Zhuping)



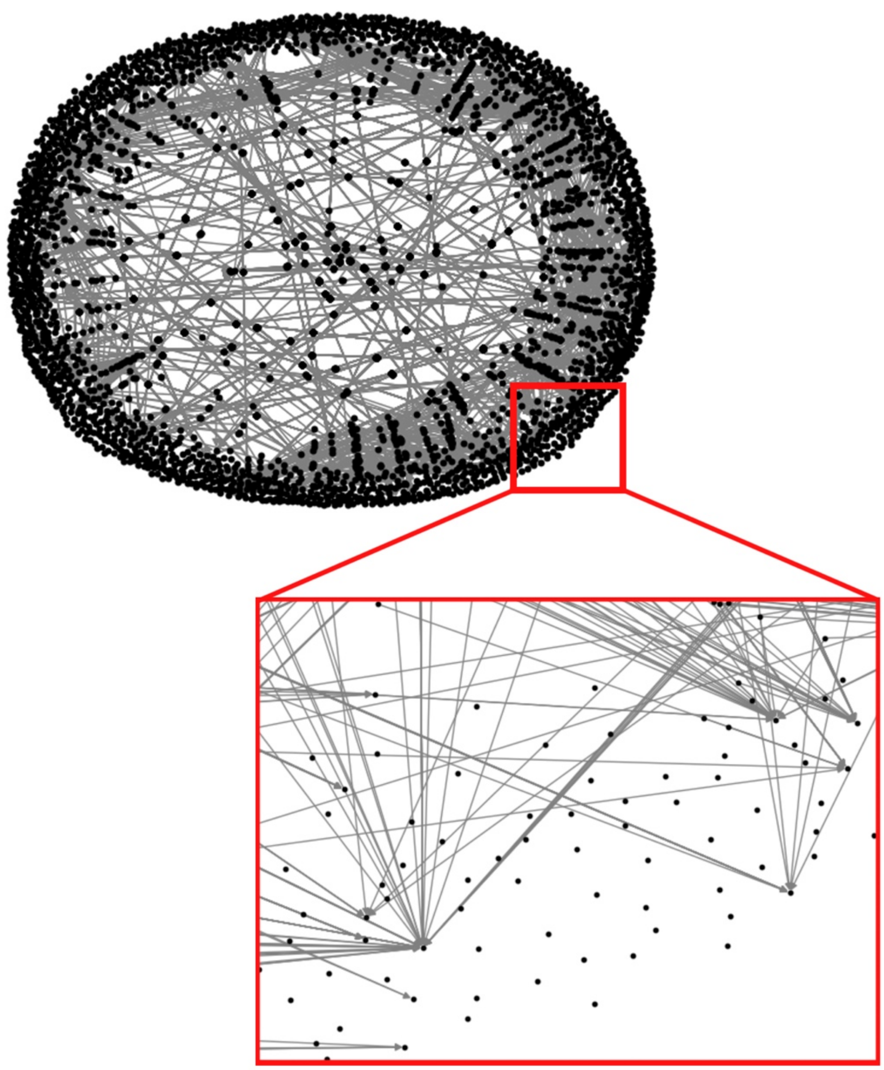
C. Promotion Decision Network

Lastly, the third graph models patronage relationship network using the information of political appointment between leaders. Specifically, I use the promotion from cadre level 4 and to level 5, i.e. deputy municipality leader to municipality heads, which is considered a milestone in one's political career. I adopt this particular point of promotion in one's career to form patronage links following Jiang (2008)'s argument:

> *"Provincial secretaries typically have an overwhelming influence over the selection of officials within provinces. Although promotions are formally decided by the provincial standing committee in a collective fashion, in practice, provincial secretaries' opinions are usually what matter the most. Organizationally, provincial secretaries are also regarded as the "person of first responsibility" (diyi zeren ren) by the higher authority when it comes to provincial-level personnel issues.*
>
> *I focus on city leadership positions because they are highly valuable posts within the Chinese system in terms of both power and upward potential. For example, 17 out of the 25 Politburo members of the 18th Central Committee had experience as city leaders earlier in their careers. Given their value and importance, it is reasonable to expect that among officials who are promoted to these posts for the first time, many will have close relations with their provincial secretaries. Even if some of them are not, we can at least assume that the party secretaries have taken a neutral stance toward their advancement, and the presence of measurement error is likely to attenuate the results toward zero."*

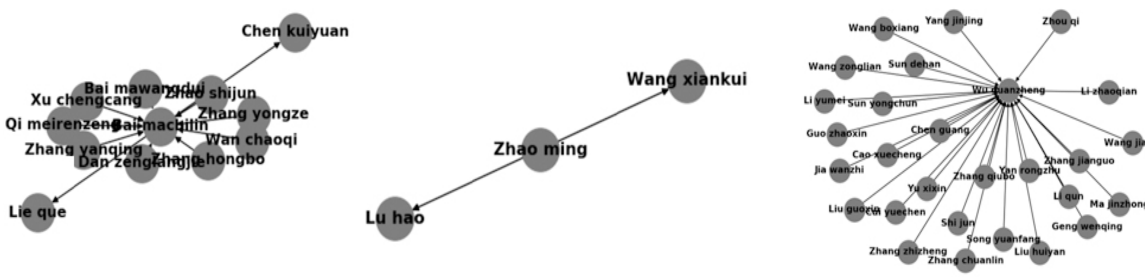
It follows that an edge between two nodes indicates a patron-client link formed through political appointment. A patron-client link from leader A to leader B forms when A is promoted to municipality head under B's leadership as party secretary or government head in the corresponding province. This network is not weighted because I am only considering one promotion for each leader. The following figure visualizes the entire appointment patronage network, with a subset of it zoomed in so as to show clear edges between nodes.

Figure 4.6: A Zoom-in Visualization of the Promotion Decision Network



Now I give an example of micro-level patronage structures in the network. To construct the following graph, I randomly select three nodes and drew out their egonets, i.e. the nodes connected to each of them. Figure 4.7 presents their patronage structures.

Figure 4.7: Egonets of Three Nodes in the Promotion Decision Network



Again, the direction of an edge (A→ B) points from a client to a patron, meaning that A is promoted to municipality head under B's leadership as party secretary or government head of the corresponding province. Therefore, a node can have at most two out-neighbors, because when one is promoted to municipality head, there can only be one provincial party secretary and one provincial head of government who are his/her promoters. The node in the left subgraph has two outgoing edges and many incoming edges, suggesting that this particular leader has promoted nearly ten clients when he held provincial head positions. Next, the node in the middle subgraph has no incoming edges, likely because she has not reached the provincial head level that would allow her to promote others. Another common structure is shown in the right subgraph; namely, the node has only incoming edges and it has a great number of them. This is a senior leader in the party, Wu Guanzheng. His patrons, i.e. leaders who promoted him, must be too old to be included in the database.

ii. Performance Data

Another important part of the data used in this paper is leaders' economic performance, particularly when they served as prefectural heads. This information is also from Junyan Jiang (2018). I quote the data discussion in his paper below.

> *"The main indicators I use to measure city leaders' economic performance are the overall and sector-specific GDP growth rates, which are collected from the Statistical Yearbooks on Regional Economy (quyu tongji nianjian) from 2000 to 2011. Although the official growth statistics are by no means free of problems (Wallace 2014), they are still the best available data to offer a consistent measure of economic performance over both time and space. To address the problem of data manipulation, I also collect alternative growth indicators that are less susceptible to official manipulation, including railway freight, power consumption, and satellite-based nighttime brightness (Henderson, Storeygard, and Weil 2012). The main sample includes all prefecture-level and subprovincial units in mainland China, except for districts under centrally administered municipalities (zhixia shi) and prefectures in Tibet. The resulting panel includes observations from 326 cities for 12 years."*

Jiang (2018) provides official and alternative growth indicators for all prefecture-level cities in Mainland China from 2000 to 2011. Combined with the information on who served as mayors and party secretaries each year, I was able to link the city-level economic performance to the leaders being studied in the patronage networks. In the following section, I will explain how I use these raw performance indicators to construct measures of leaders' genuine ability in generating growth, thereby overcoming the potential selection bias from omitting patronage information.

V. Measurement Strategy for Patronage and Performance

Having introduced patronage network graphs and performance data in the last section, I explain how I use them to build comprehensive measures for statistical inference. In other words, this section shows the construction of the right-hand side variables that will later go into my prediction models. In particular, I adopt a new network-based measurement strategy to capture neighborhood features of leaders in the patronage networks.

i. Measuring Patronage Networks

While the previously built patronage graphs demonstrate connections visually, to use leaders' networks for quantitative analysis requires both the extraction and representation of graphical links surround each leader in the network. Substantively, we need a way to quantify what the neighborhoods of each leader in the network look like, so as to capture the patronage information needed for prediction models. I discuss two methods of doing so and compare their effectiveness in the prediction results part in the next section. Both methods quantify neighborhood views for leaders in patronage networks, but they do so in technically different ways. It turns out that the second method, node2vec (*Random Walks)*, which adopts a state-of-the-art computer science feature learning algorithm, works much better than the alternative.

Method 1: Two-Hop Neighborhood Features

The first is to generate two-hop neighborhood features for each node in the graph as follows. I start by choosing 3 basic local features for each node (in this order):

1. The degree of v, i.e. the number of its neighboring nodes;

2. The number of edges in the egonet of *v*, where the egonet of *v* is defined as a subgraph whose nodes are *v*, and its neighbors and edges are induced from the whole graph;
3. The number of edges that connects the egonet of *v* with the rest of the graph, i.e. the number of edges that enters or leaves the egonet of *v*.

Intuitively, these three basic features at the node level measures the number of leaders one connects to, the connectivity strength between one's neighboring leaders, and the strength between one's neighborhood and the rest of the entire network. The next step is to recursively generate more features using the mean and sum as aggregation functions.

Initially, we have a feature vector $V_u \in R^3$ for every node u. In the first iteration, I concatenate the mean of all u's neighbors' feature vectors to V_u, and do the same for sum, i.e.

$$\tilde{V}_u^{(1)} = [\tilde{V}_u; \frac{1}{|N(u)|} \sum_{v \in N(u)} \tilde{V}_v; \sum_{v \in N(u)} \tilde{V}_v] \in \mathbb{R}^9,$$

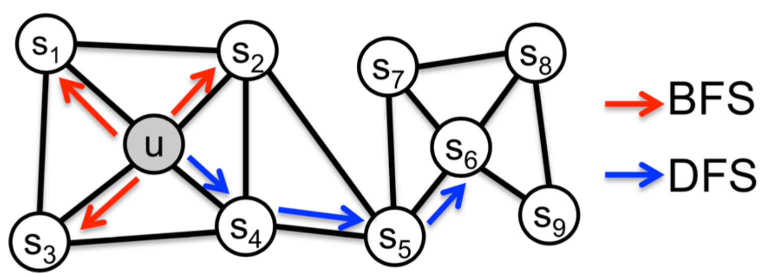
which gives me a one-hop feature vector of length 9. Repeating it through a second iteration, I obtain a two-hop feature vector of length 27 for each node. In other words, I have for each leader a 27-dimensional measure of the size, strength, and connectivity of the connections for herself, her neighboring nodes, and the neighbors of her neighboring nodes.

This method gives us a fixed representation of local network structures surrounding a node through examining her/his neighborhoods of different distances.

Method 2: Node2Vec (*Random Walks*)

The second method is to extract and transform neighborhood features into vectors using *node2vec*, which works by carrying out a number of random walks from each node in the graph. Here, the walks are parameterized by p and q (Grover and Leskovec 2016). This method can do BFS (Breath-First Sampling) to learn structural equivalences between nodes by looking at immediate neighbors of the source node. For example, structural equivalence based on network roles such bridges and hubs can be inferred just by observing the neighborhood of every node. Besides BFS, it can also conduct DFS (Depth-First Search) to examine nodes sequentially sampled at increasing distances from the source node, which explores larger parts of the network as it can move further away from the source node. The sampled nodes more accurately reflect a macro-view of the neighborhood that is essential for inferring communities based on homophily as oppose to the structural equivalences given by BFS.

Figure 5.1: BFS and DFS search strategies from node u (Grover and Leskovec 2016)



The node2vec algorithm developed in Grover and Leskovec (2016) is a state-of-the-art algorithmic framework that can carry out both BFS-like and DFS-like searches for learning continuous feature representations for nodes in networks. In my study, I use node2vec to generate two sets of 128-dimension feature vectors (one DFS, the other BFS) for each node in order to capture the micro- and macro- views of its extended neighborhoods.

In sum, this *node2vec* method gives us numerical representation of both the local neighborhoods (micro-views) and global networks (macro-views) for each leader in the entire nationwide patronage graph. Although still relying on one-to-one identification of political connection, this new measurement and representation strategy quantifies what local patronage networks for a leader look like, with a flexible notion of neighborhoods. From the above two methods, I obtain two sets of patronage features for each leader: 1) a 27-dimension two-hop neighbor representation, and 2) a 256-dimension BFS and DFS combined feature vector from node2vec (*Random Walks)*. What follows in this section is a discussion of how this paper constructed absolute and relative performance measures for each leader under investigation.

ii. Measuring Performance

The following part lists out the two sets of performance measures I use for prediction purposes. Each data point is a performance indicator for a leader averaged over his/her years serving as a prefectural head (mayor or party secretary). The first set contains absolute measures of economic growth during the time a leader serves as head of a prefecture city, while the second set controls for starting GDP size, past and subsequent performance of the city, and common shocks to serve as relative measures of performance.

Absolute City-Leader Level Performance Variables[1]:

1. *GDP growth rate*
2. *GDP growth rate in agriculture*
3. *GDP growth rate in industry*
4. *GDP growth rate in service*
5. *GDP growth rate at year + 2*
6. *Log average nighttime brightness per square km*

[1] The first measure, the overall GDP growth rate, is a weighted average of the second, third, and fourth measures, GDP growth rates in agriculture, industry, and services, with weights given by the level sizes of these sectors.

Relative Performance Measures:

Each of the absolute performance variables above, but after controlling for

1. *City fixed effects*
2. *Province-year fixed effects*
3. *GDP size and growth at the starting year of the current prefectural heads*

The residual variables after controlling are relative measures of a city's economic growth when comparing it to how it is doing in the past and in the subsequent years (city fixed effects), to how other surrounding cities are doing in the same year (province-year fixed effects), and to how it is when current leaders first come into office. In doing so, it removes the selection bias that the leaders with stronger patronage networks are placed *ex ante* in places more likely to grow faster. We can therefore regard this set of relative performance measures as a proxy for the genuine ability of leaders in generating economic growth. Besides performance and patronage measures, I also collect a set of covariates at the leader-level:

1. *Gender*
2. *Ethnicity*
3. *Age as of 2005*
4. *Education level as of 2005*
5. *Years of party membership as of 2005*
6. *Rank level in 2005*

In the next section, I show results from several types of model with various different specifications to compare the relative importance of patronage and performance in determining future promotion.

VI. Predictive Models and Results

In this section, I use four types of models, from the most commonly used linear regressions to the advanced complex neural networks, to predict cadre rank outcome in 2015, based on performance measures and patronage information extracted in and prior to 2005. By holding off information post-2005, this empirical exercise serves to test the relative prediction power of patronage and performance in affecting future political promotion.

This section is divided into three parts. I firstly give a brief discussion on the four models I choose for prediction, i.e. *Ordinary Least Squares, Ordered Logistic Regression, Random Forests, and Neural Networks*. I explain their assumptions, advantages, limits, and how they differ and relate to each other conceptually. Next, I plug in patronage variables into each of the models under different model specifications to test their predictive power. Lastly, I show the power of performance measures when used alone and when combined with patronage information, which allows for comparing the relative importance of patronage and performance in driving promotion.

i. Model Choices

OLS: Ordinary Least Squares regression is the most commonly used model in econometric analysis. It is easy to estimate and computationally cheap. Most importantly, it imposes a strong functional form assumption that relationships between the outcome variable Y and the predictors X (right-hand side variables) are linear. The advantage of linearity assumption is high interpretability – model specification is straightforward and parameter estimates make intuitive sense. The downside, however, is that it fails to capture nonlinear, multi-stage relationships among predictors and between predictors and the outcome variable. Therefore, it is usually included as a standard baseline model to compare with other advanced models. The following equation gives the functional form of OLS model with full specification of features.

$$Final\_Rank = \alpha + \beta V_{patronage\_features} + \gamma V_{performance\_features} + \delta C_{covariates} + \epsilon$$

As shown in the last section, the set of patronage features is a 256-dimension vector from *Random Walks* that capture patronage neighborhood information, the set of performance features include *GDP growth rate, GDP growth in agriculture, GDP growth in industry, GDP growth in service, GDP growth rate at year + 2*, and logged average nighttime brightness per square km (in either their absolute or relative forms). The covariates set consists of *Gender, Ethnicity, Age as of 2005, Education level as of 2005, Years of party membership as of 2005, Rank level in 2005.* I will present regression results in table in the later estimation part.

Ordered Logistic Regression: This ordered logit model, a.k.a. proportional odds logistic regression, is particularly appropriate here, because our Y variable, rank outcome, is an ordered categorical response. It is better than continuous outcome models like OLS, for it does not assume equal distances between levels; it also preserves ordering information in the outcome variable compared to categorical outcome models like multinomial logistic regression. The model is set up as follows. We first obtain a latent variable representation Y_i^* using the exact same formula as in OLS:

$$Y_i^{\star} = \alpha + \beta V_{i,patronage\_features} + \gamma V_{i,performance\_features} + \delta C_{i,covariates} + \epsilon_i$$

Then we map the latent strength Y_i^* to Y_i using the following scheme:

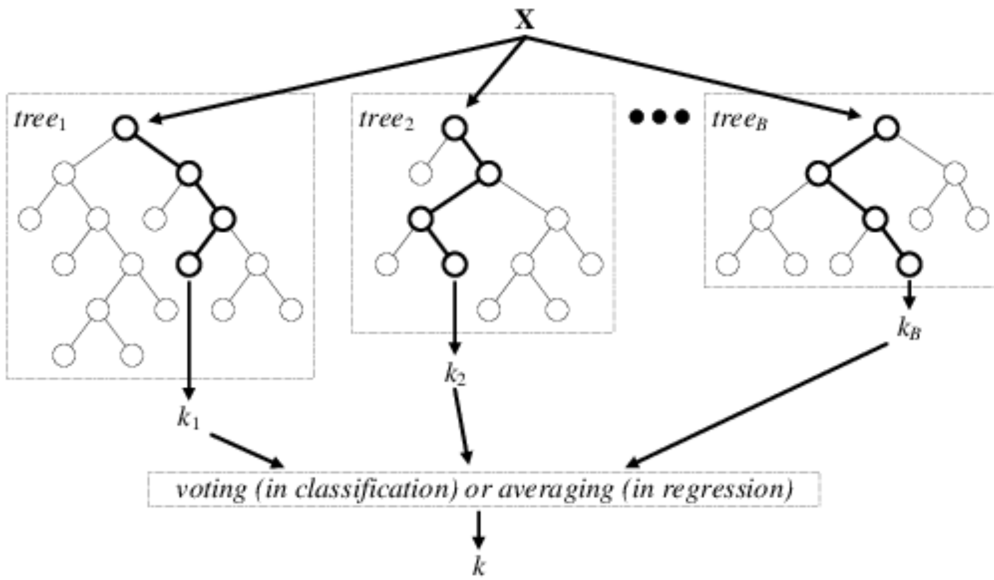
$$Y_i = \begin{cases} 1 & \text{if} \quad -\infty(=\psi_0) < Y_i^* \leq \psi_1, \\ 2 & \text{if} \quad \psi_1 < Y_i^* \leq \psi_2, \\ \vdots & \vdots \\ J & \text{if} \quad \psi_{J-1} < Y_i^* \leq \infty(=\psi_J) \end{cases}$$

where these φ's are threshold parameters to be estimated. Lastly, assuming the errors are *i.i.d.* logistic, we derive probabilities of Y given our X data using estimated φ parameter values as follows:

$$\Pr(Y_i \leq j \mid X_i) = \frac{\exp(\psi_j - X_i' \beta)}{1 + \exp(\psi_j - X_i^\top \beta)}$$

Here Y_i is the final rank outcome for leader *i*. The estimation of this ordered logit model is relatively easy, but interpretation breaks down as it assumes a latent representation. Therefore, parameter estimates cannot be read directly as those from OLS for the direction and magnitude of correlations between predictors and outcome. Although having added a stochastic component, this ordered logit model still assume a strong linear functional form in its latent variable as that in OLS. Both OLS and ordered logit are fully parametric methods and serve as basic tools in econometrics. Next we move to more advanced methods from machine learning, which are nonparametric and allow for complex, nonlinear relationships between predictors and the outcome variable.

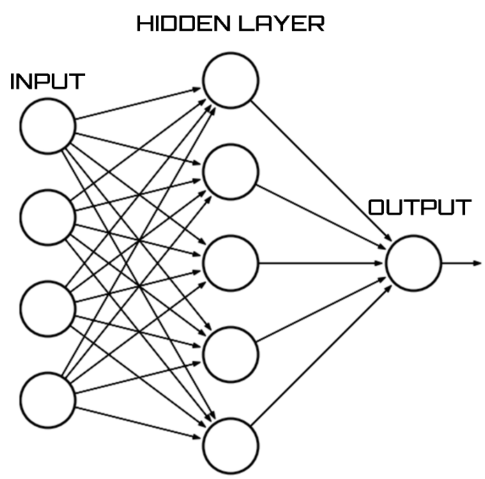
Random Forests: Since we are faced with a classification task in which we need to categorize rank outcomes into several different levels, the Random Forests (RFs) classifier comes in as a natural choice in machine learning. The RF model works by bootstrapping the original dataset with replacement to build thousands of single decision trees and then aggregates results across all the trees. A single decision tree is grown by recursively partitioning the feature space of right-hand side variables (i.e. predictors) to learn the classification output of the response variable. The following graph[1] provides some intuition about the process.



Without going into further detail of the learning and estimation process, two key things need to be pointed out. First, the RFs model is a fully nonparametric method that imposes no functional form assumption, which means it allows for complex interactions and non-linear relationships; however, this also implies that we lose the interpretability of model process and parameter estimates. Secondly, RFs are proved to work well particularly when there are few observations and many right-hand side predicting variables. A full list of input variables includes patronage measures, performance measures, and covariates as discussed earlier, while I will indicate specific specification when giving each result in the later prediction part.

[1] Source: Verikas, Antanas & Vaiciukynas, Evaldas & Gelzinis, Adas & Parker, James & Olsson, M. Charlotte. (2016). Electromyographic Patterns during Golf Swing: Activation Sequence Profiling and Prediction of Shot Effectiveness. Sensors. 16. 592. 10.3390/s16040592.

Neural Network: Lastly, I introduce the neural network model, which is particularly relevant here not because it is currently a cutting-edge model in the field of machine learning and artificial intelligence, but because its modeling structure can potentially work well with the patronage vector representation from our new network-based measurement strategy using *Random Walks*. The whole idea of artificial neural network is based on the concept of the structure and functions of a human brain. A human brain consists of neurons that process and transmit information between themselves. There are dendrites that receive inputs and produce outputs through an axon to other neurons. The imitation of the human brain in machine learning allows for advanced statistical modeling structures that are able to capture complex interactions and non-linear relationships between variables of interest. The graph below is a toy example of the modeling structure of the neural network model.



Particularly, the input layer neurons feed in a 256-dimension vector embedding from node2vec (*Random Walks)* which carries patronage networks information prior to 2005 for each leader. The output layer is a log-logistic multinomial function for predicting cadre ranks in 2015, and there are in total 260 neurons in the single hidden layer in between. This NNs model can also feed in any other RHS variables, including economic performance measures, which I will test and compare with patronage variables in the following part. While also a nonparametric machine learning model, NNs do not directly relate to RFs in concept. The essence of RFs is to try different ways of feature space splitting and aggregate single trials to output a forest-like decision. In contrast, NNs are most concerned with changes in the flow of information through neurons, include their existence, direction, and strength, which is highly similar to the setup and structure of nodes and edges in network analysis. This, then, gives us a reason to expect superior performance when NNs and measures from patronage networks are combined.

ii. Predicting with Patronage

In this part, I predict political promotion using patronage networks information alone across four types of model discussed above. I test with patronage measures alone first to check whether patronage information is an important predictor of future promotion and also to select a better patronage measurement (*two-hop* features or *Random Walks*) for subsequent use. In each of the four model classes, i.e. *OLS, Ordered Logistic Regression, Random Forests, and Neural Networks*, I try to estimate with following four specifications:

1. *An intercept only (null model)*
2. *Patronage two-hop neighborhood features*
3. *Patronage Random Walks features*
4. *All features (two-hop + random walks + covariates).*

The following table summarizes the goodness of fit across different model class choices and different model specifications, each evaluated by R^2 (percent of variation explained in outcome variable), Accuracy (percent of correctly predicted cadre ranks), MSE (mean squared errors of predicted ranks), and Residual Deviance. The "Predictors" column specifies the RHS variables.

Table 6.1: Goodness of Fit for Patronage Models

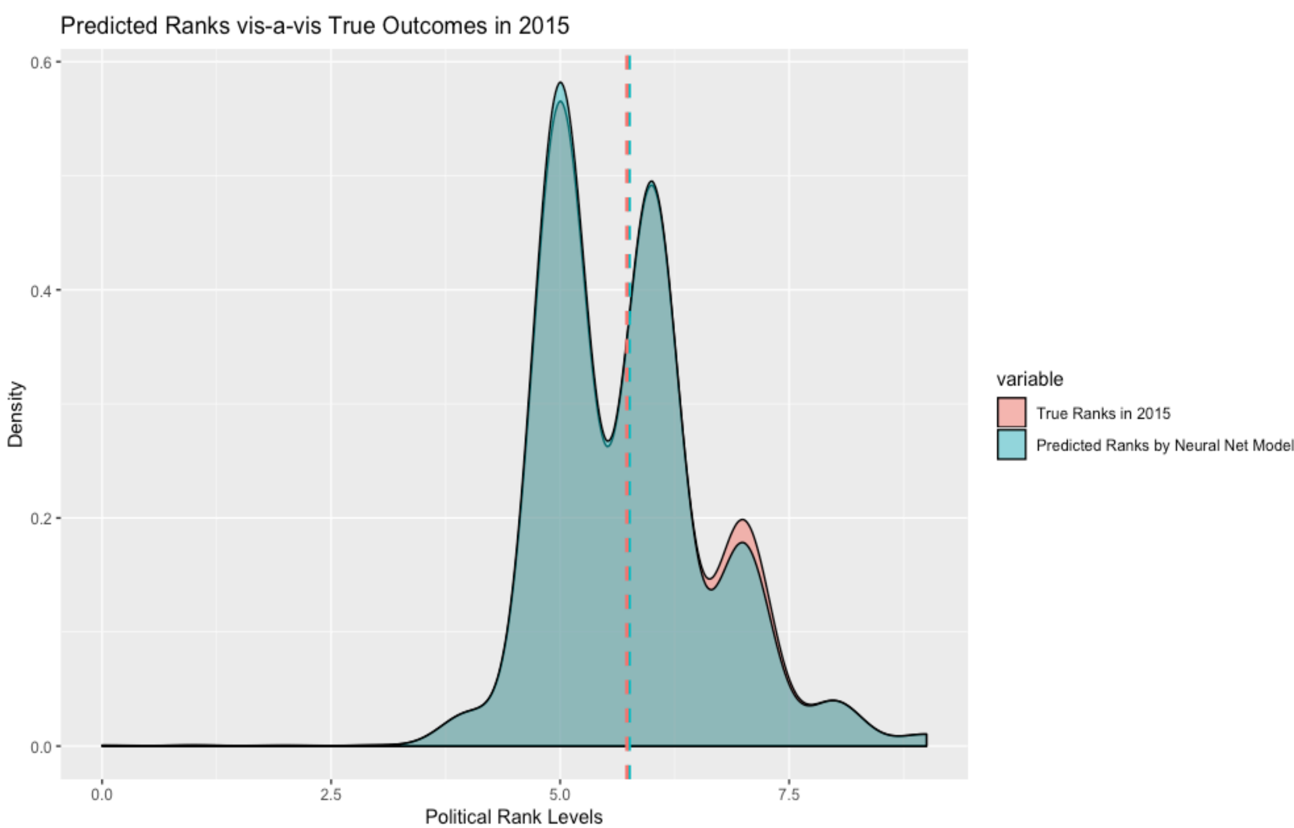
	Predictors	OLS	Ordered Logit	Random Forests	Neural Networks
Percent of Variation Explained in Rank Outcomes	No Feature	0.00	–	–	–
	Two-hop Neighbors	67.79	29.72	71.42	12.40
	Random Walks	70.24	–	71.26	**78.43**
	All Features	71.66	–	72.59	16.18
Percent of Correctly Predicted Cadre Ranks	No Feature	36.68	42.60	–	42.60
	Two-hop Neighbors	66.04	65.74	70.92	53.39
	Random Walks	69.17	–	72.83	**82.74**
	All Features	71.19	–	73.42	55.18
Mean Squared Errors of Predicted Ranks	No Feature	0.82	1.39	–	1.39
	Two-hop Neighbors	0.27	0.58	0.24	0.72
	Random Walks	0.25	–	0.24	**0.18**
	All Features	0.23	–	0.23	0.69
Residual Deviance In the Model	No Feature	–	8441.59	–	8441.59
	Two-hop Neighbors	–	5468.02	–	6848.22
	Random Walks	–	–	–	**2896.62**
	All Features	–	–	–	6740.65

The above table reports results of estimating each of the four types of model on each of my feature sets (model specification) and evaluated by four major measure of goodness of model fit. The "—" indicates either a measure is not applicable to the model or where model cannot converge.[1] The bolded numbers in the table refer to the best performance achieved with regard to each of the four evaluation criteria. It is remarkable that the best performances resulted from a single-layer neural networks model on the 256-dimension patronage network vector representation learned by the *Random Walks* feature set.

[1] For example, there is no residual deviance measure for OLS or RFs. Also, ordered logit failed to converge using Random Walks set and all feature set for there were too many RHS variables.

This choice of modeling and feature set achieved an R^2 of 78% and an accuracy of 83%. In other words, using merely information embedded in patronage networks until 2005, the model is able to explain, ten years later, nearly 80 percent of the variation in cadre rank outcomes, and predict rank outcomes with 83 percent of accuracy. The predictors used here even do not include the covariate set, e.g. leader gender, age, ethnicity, party experience, etc., but only come from vector representations of patronage network structures and homophiles constructed ten years prior to the outcome. The following figure presents the predictions from the neural network model vis-à-vis the true outcome of cadre ranks in 2015. We can see that the predicted ranks capture the true outcomes almost perfectly in most parts of its distribution. The only exception is around rank level 7 where the model predictions were off by a small margin.

Figure 6.1: Comparison Between Predictions and True Outcomes



iii. Predicting with Performance

While patronage serves as a very strong predictor of elite advancement as shown above, we need to compare it with performance measures in order to test our hypotheses. As the patronage features from *Random Walks* perform much better than the two-hop neighborhood features, I will use *Random Walks* representation as the patronage features going forward.

Again, we start with our baseline model: linear regression model. The following table presents OLS regression results when estimating separately using patronage and performance and when the two metrics are combined, using both relative and absolute performance measures. Model 1 uses only patronage measures, Model 2 uses only relative performance measures, Model 3 uses only absolute performance measures, and Model 4 and 5 combine patronage with relative and absolute performances. Because patronage information is a long vector of length 256 from Random Walks, I only report those coefficients that are statistically significant at the 5% level, i.e. having at least two stars.

Two important observations are evident in the OLS results table. First, patronage variables emerge much more statistically significant than performance variables, suggesting a much greater predictive power on the outcome variable. Although there are nearly twenty patronage variables with at least two stars, only one performance variable (the absolute GDP growth rate of the year) came out significant. Second, models with patronage variables always perform better than those without them (Models 1, 4, 5 vis-à-vis Models 2, 3) in both Accuracy (percent of correctly predicted cadre ranks) and R^2 (Percent of variation explained in rank outcomes) model evaluation metrics.

OLS Results for Patronage and Performance Comparison

	Dependent variable: rank_2015				
	Model 1	Model 2	Model 3	Model 4	Model 5
	Patronage Networks	Relative Performance	Absolute Performance	Patronage + Performance(rlt)	Patronage + Performance(abs)
gdpidx_rl		0.007		0.004	
		(0.016)		(0.016)	
gdpidx_1st_rl		−0.004		−0.002	
		(0.006)		(0.005)	
gdpidx_2nd_rl		−0.005		−0.002	
		(0.007)		(0.007)	
gdpidx_3rd_rl		−0.001		−0.003	
		(0.010)		(0.010)	
f2gdpidx_rl		0.018		0.014	
		(0.011)		(0.011)	
logltavg_rl		0.131		0.182	
		(0.129)		(0.125)	
gdpidx			0.022**		0.013
			(0.009)		(0.009)
gdpidx_1st			−0.006		−0.006
			(0.004)		(0.004)
gdpidx_2nd			−0.004		−0.002
			(0.005)		(0.005)
gdpidx_3rd			0.008		0.006
			(0.006)		(0.006)
f2gdpidx			0.0005		0.004
			(0.007)		(0.007)
logltavg			0.005		0.015
			(0.010)		(0.011)
‘rw-d4‘	−3.288**			−3.289**	−3.150**
	(1.518)			(1.519)	(1.517)
‘rw-d7‘	3.376**			3.353**	3.470**
	(1.408)			(1.409)	(1.406)
‘rw-d9‘	2.715**			2.662*	2.686**
	(1.360)			(1.363)	(1.357)
‘rw-d23‘	3.663**			3.647**	3.764**
	(1.582)			(1.583)	(1.579)
‘rw-d27‘	3.723**			3.814***	3.870***
	(1.449)			(1.453)	(1.448)
‘rw-d45‘	3.225**			3.200**	3.350**
	(1.374)			(1.376)	(1.373)
‘rw-d48‘	3.281**			3.310**	3.328**
	(1.565)			(1.567)	(1.563)
‘rw-d93‘	3.473**			3.518**	3.528**
	(1.427)			(1.428)	(1.425)
‘rw-d119‘	4.439***			4.485***	4.512***
	(1.429)			(1.431)	(1.427)
‘rw-d142‘	−2.445**			−2.466**	−2.634**
	(1.204)			(1.206)	(1.203)

OLS Results for Patronage and Performance Comparison (Cont'd)

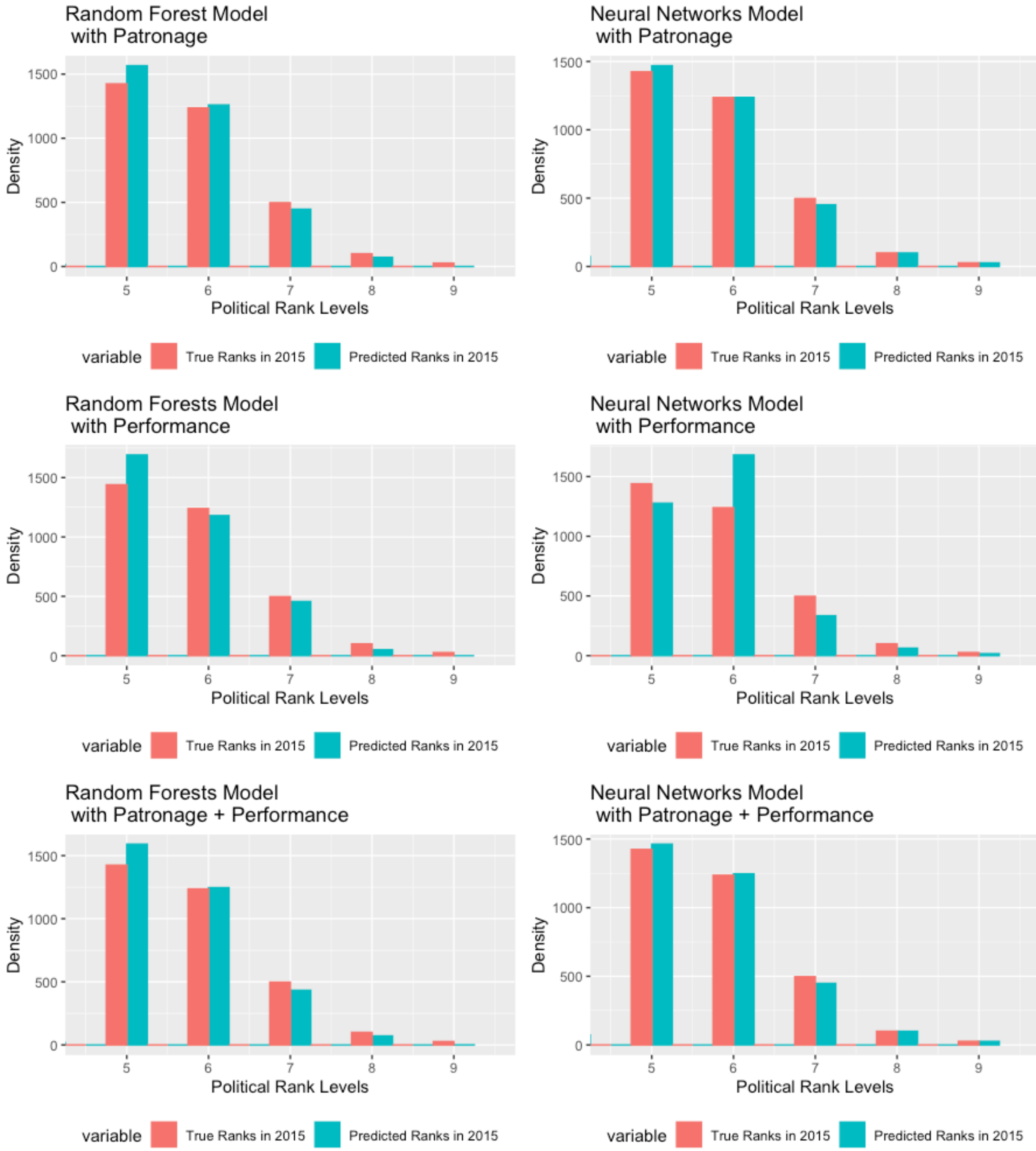
	Dependent variable: rank_2015				
	Model 1	Model 2	Model 3	Model 4	Model 5
	Patronage Networks	Relative Performance	Absolute Performance	Patronage + Performance(rlt)	Patronage + Performance(abs)
'rw-d147'	2.529**			2.474**	2.564**
	(1.090)			(1.092)	(1.089)
'rw-d171'	−3.315***			−3.297***	−3.131***
	(1.130)			(1.131)	(1.129)
'rw-d193'	−3.350**			−3.404**	−3.435**
	(1.460)			(1.462)	(1.459)
'rw-d208'	−3.768***			−3.748***	−3.624***
	(1.229)			(1.229)	(1.227)
'rw-d223'	3.325**			3.343**	3.362**
	(1.349)			(1.349)	(1.346)
'rw-d224'	2.995**			2.989**	2.956**
	(1.179)			(1.181)	(1.177)
'rw-d230'	2.362**			2.429**	2.392**
	(1.137)			(1.137)	(1.134)
'rw-d234'	−2.906**			−2.953**	−2.771**
	(1.381)			(1.382)	(1.379)
'rw-d244'	3.095**			3.108**	2.980**
	(1.389)			(1.391)	(1.387)
rank_2005	0.554***	0.603***	0.604***	0.555***	0.555***
	(0.010)	(0.008)	(0.008)	(0.010)	(0.010)
Constant	2.967***	2.652***	2.378***	2.969***	2.719***
	(0.198)	(0.042)	(0.082)	(0.198)	(0.211)
Observations	3,360	3,378	3,378	3,360	3,360
R^2	0.702	0.633	0.636	0.703	0.704
Accuracy (%)	**69.17**	**63.31**	**61.28**	**69.23**	**69.79**

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Taken together, this table tells us that even using the baseline OLS model, patronage information has greater predictive power than economic performance, and the marginal gains from including performance measures are minimal once we have patronage information. What follows is estimation through other advanced models to examine whether this pattern holds and to check predictive power across models and feature sets.

Since parameter estimates in the Random Forests and Neural Networks models cannot be interpreted directly, I show their prediction results vis-à-vis true outcome as follows.

Figure 6.2: Comparison Between Predicted and True Outcomes: RFs and NNs Models



The left three plots are from Random Forests, while the right-hand ones are from Neural Networks. Each group has three model specifications: Patronage based measures only, Performance based measures only, and combining Patronage and Performance based measures. Two observations are crucial here: first and most importantly, within both the RFs group and the

NNs group, patronage based measures always do better than performance-based measures (comparing Row 1 with Row 2 models), and the marginal gains from further including performance based measures are minimal once we have patronage measures (comparing Row 1 and Row 3 models). In particular, the results from NNs models are almost identical between only using patronage information and combining both patronage and performance, suggesting that performance measures do not give any additional benefit of prediction when we have NNs models working with patronage information. Second, the best-performing models are from NNs, which defeat RFs and do an almost perfect job in predicting the highest rank levels (Level 8 and Level 9). NNs models also almost eliminate the large prediction errors RF models have for Level 5 officials, with only 45 incorrect predictions there in the Patronage only specification (1469 predicted vis-à-vis 1424 true ranks in Level 5) and less than 40 incorrect predictions there in the Patronage + Performance specification (1463 predicted vis-à-vis 1424 true ranks).

Now let us further investigate the relative importance of performance-based and patronage-based measures in the RFs model. The following table shows the increase of node purity and the percent increase in out-of-bag MSE for each variable as a result of this variable being permuted, that is, of values being randomly shuffled. In other words, these are indicators of how much *worse* a model performs had it not included (the real values of) a variable. Therefore, both *IncNodePurity* and *%IncMSE* measure the importance of variables in a fitted RFs model, with higher values corresponding to higher importance.

Table 6.1b: Variable Importance in the RFs Model

Patronage Vars	%IncMSE	IncNodePurity	Performance Vars	%IncMSE	IncNodePurity
rw_d230	13.78	45.72	*gdpidx_2nd_rl*	9.74	11.94
rw_d94	13.27	16.46	*gdpidx_3rd_rl*	9.47	13.06
rw_d246	13.10	9.20	*logltavg_rl*	8.56	12.71
rw_d146	12.12	17.96	*gdpidx_rl*	8.42	10.20
rw_d175	11.82	53.36	*f2gdpidx_rl*	7.26	5.54
rw_d132	11.22	11.97	*gdpidx_1st_rl*	7.26	10.42

The right-hand part reports importance results for all of the six performance variables used in the RFs model, while the left-side reports the top six patronage variables, both in a decreasing order. It is obvious that patronage variables have higher *%IncMSE* and higher *IncNodePurity* than those of performance variables. Even the sixth patronage variable, *rw_d132,* outperforms the top performance variable, *gdpidx_2nd_rl,* in both measures. Recall that a high value suggests a larger loss of model performance if the variable is excluded. Therefore, the above results overwhelmingly point to the importance of patronage-based measures in relative to performance when predicting promotion. Lastly, let us summarize results from all of the OLS, RFs, and NNs models in a table.

Table 6.2a: Comparing Predictive Power between Patronage Measures and Relative Performance Measures

		Patronage	Relative Performance	Patronage + Relative Performance	Marginal Gains
Percent of Variation Explained in Rank Outcomes	OLS	70.24	61.07	70.28	0.04
	Random Forests	71.37	69.82	72.07	0.70
	Neural Networks	**78.43**	58.54	78.16	**-0.27**
Percent of Correctly Predicted Cadre Ranks	OLS	69.17	63.31	69.23	0.06
	Random Forests	72.59	69.63	73.57	0.98
	Neural Networks	**82.74**	68.83	82.50	**-0.24**
Mean Squared Errors of Predicted Ranks	OLS	0.25	0.30	0.25	0.00
	Random Forests	0.24	0.25	0.23	0.01
	Neural Networks	**0.18**	0.34	0.18	**0.00**

The above table summarizes the comparison between patronage networks and relative economic performance regarding their predictive power on political promotion across all models.[1] Column 1 (Patronage) gives results from using only patronage information in OLS, Random Forests, and Neural Networks models. Column 2 (Relative Performance) shows results from using only relative measures of economic performance. Column 3 (Patronage + Relative Performance) gives outcomes when patronage information is combined with relative economic performance of the leaders. And Column 4 (Marginal Gains) shows the difference between before and after including relative economic performance measures with patronage information, i.e. gap between Column 3 and Column 1.

This table tells us two important things. First, when using patronage or performance measures separately, patronage defeats performance in all three model evaluation criteria across all three classes of model. Particularly in the Neural Networks model,

[1] The ordered logit model class was dropped because it did not run or converge with majority of our data structures and whenever it worked, it gave the worst performance as shown in Table 6.1.

patronage information explains nearly 20 percentage points *more* variation in career outcome and has an accuracy about 14 percentage points *higher*, suggesting patronage is a much stronger predictor than economic performance. Secondly, further including performance measures does not give us a better prediction of cadre rank outcomes. As shown in Column 4, the gains in OLS and Random Forests models were minimal – all less than one percentage point. Furthermore, it yields a worse result when predicting in our Neural Network model, which is the best-performing model in our study. The facts that marginal gains from including performance measures in patronage information are either neglectable or negative suggests that performance measures do not contain additional information regarding promotion once we have patronage networks. The following table shows the results from using absolute measures of economic performance, which point to the same conclusions.

Table 6.2b: Comparing Predictive Power between Patronage Measures and Absolute Performance Measures

		Patronage	Absolute Performance	Patronage + Absolute Performance	Marginal Gains
Percent of Variation Explained in Rank Outcomes	OLS	70.24	63.58	70.42	0.18
	Random Forests	71.37	71.21	72.17	0.80
	Neural Networks	**78.43**	62.82	78.29	**-0.14**
Percent of Correctly Predicted Cadre Ranks	OLS	69.17	61.28	69.79	0.62
	Random Forests	72.59	71.52	73.51	0.92
	Neural Networks	**82.74**	72.05	82.62	**-0.12**
Mean Squared Errors of Predicted Ranks	OLS	0.25	0.30	0.25	0.00
	Random Forests	0.24	0.24	0.23	0.01
	Neural Networks	**0.18**	0.31	0.18	**0.00**

Again, we see that patronage always does much better than performance, and when using patronage and performance together, the results get worse. This again indicates that patronage remains the best predictor of promotion even after performance is controlled for.

To sum up, in this section, I first show the construction of patronage and economic performance (absolute and relative) measures. Then I use patronage and performance measures separately and together across different models to test their predictive power on future elite advancement. No matter which model class or what model evaluation criterion is used, patronage information always performs better than the economic performance measures, even after controlling for performance. Furthermore, the striking fact that adding performance does not improve the patronage predictions suggests that performance measures contain no additional information of promotion prospects other than those in patronage networks.

The results overwhelmingly favor our second hypothesis that patronage, rather than performance, dominantly drives the selection of leaders. In fact, the statistical patterns found in this section shows that once patronage information is taken into account, the predictive power of economic performance on career advancement entirely goes away.

VII. Conclusion and Discussion

In this paper, I demonstrate that political patronage, rather than economic performance, is the dominant driver of career advancement for middle-level leaders and above in contemporary China. I do this by drawing from the latest systematic data on leaders' career and performance data and applying advanced network analysis and machine learning methods. In particular, I show that information encoded in patronage networks formed early in their career alone predicts future promotion with high accuracies. Not only do economic performance measures alone perform far worse in predicting promotion, but, once we have patronage information, performance data tells us nothing more about promotion.

Implication

There are two important implications of this study. First, the simple fact that we are able to use the information encoded in patronage networks to accurately predict ten years later what happens to officials' careers regarding their rank levels serves as a strong piece of evidence for early determination of elite career advancement. A leader's early interactions with other leaders in networks, formed into patronage links, structures, and homophiles, carry significant information about his/her political success or failure many years later. This suggests that a political leader's career outcome may have been determined much earlier than the literature think. The majority of extant literature on the determinants of elite ranking and promotion in China focuses on leaders at the provincial-level and above, i.e. elite ranking within the Central Committee (e.g. Shih, Adoph, and Liu 2012) or promotion patterns for provincial heads (e.g. Li and Zhou 2005). This study shows that what leaders have done early in their careers matter a great deal. Therefore, instead of looking at factions among top elites and provincial leaders, scholars can benefit more from moving downward to examine prefectural leaders, since the final outcome of their career may have already been determined by what their patronage networks looks like in the early stages of their career.

Secondly, this paper advances a network-based perspective of understanding and examining patronage relationships. Put differently, it advocates for a move from one-on-one patron-client links to many-to-many clientelist networks. A leader can have multiple patrons with heterogenous strengths, and at the same time, serves as a patron to multiple clients. Such a perspective asks scholars to think and study patronage in groups and further explore the group effect of patronage on individual career outcomes. In particular, the new measurement strategy – through *Random Walks* – adopted in the paper enables quantitative representation of patronage neighborhoods for leaders in the networks, which contain useful information about leaders' complex patronage relationships and has been shown to be highly predictive for individual career tracks and also for the evolution of elite networks as a whole. Moreover, the combination of network analysis methods to construct patronage network graphs and the use of a deep learning neural network model yields good prediction results, suggesting these new methods can work well with large-sized data of complex structures (we have about 0.6 million edges in complex networks). Additionally, they can bring to political scientists new measurement strategies and powerful tools for novel statistical inference.

Limits and Future Work

There are a number of major limitations to this study and a couple of corresponding unanswered questions and tasks left for future work. In the data and methods part, first, the sample spans over only twenty years in a single country. To improve internal validity, one can collect more data on China and test with year-to-year multiple rounds using prediction windows of varying sizes. Before generalizing this study to other authoritarian regimes, a standardized prototype is first needed to assist future researchers to test patronage, performance, and promotion in heterogenous bureaucratic contexts. Secondly, this study has not taken into account the strength and direction of patronage links. All that has been used is the existence of patronage edges in networks. Also, for the outcome of interest, it does not consider prosecution (a leader taken down by corruption investigation) besides looking at changes in cadre rank. Future work can incorporate new data, refine current measurements, and further explore statistical methods to include information on patronage strength and direction.

In theory part, we still do not know the substantive meanings that patronage networks determine political promotion. There remain significant room for theoretical exploration. Future work can probe into the specific structural and homophily patterns of patronage networks and compare those who climb up quickly with those who do so slowly. One can also show how some subsets of patronage networks move upward together over time and discuss possible patterns and learnings there. Moreover, while this paper studies the network effect of patronage on promotion, it does not tell us about the formation of patronage networks in the first place. How do leaders establish or enter certain types of networks and not others? Relatedly, how does a patron or a patronage network recruit their clients? What determines admission into a patronage network with promising career prospects? These unanswered questions become particularly important when we adopt an expanded, network-based view of patron-client relationship to study its origins and consequences.

References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. "Coalition formation in non-democracies." *The Review of Economic Studies* 75, no. 4 (2008): 987-1009.

Boix, Carles, and Milan W. Svolik. "The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships." *The Journal of Politics* 75, no. 2 (2013): 300-316.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson, and James D. Morrow. 2003. *The Logic of Political Survival.* Cam- bridge, MA: MIT Press.

Burns, J.P., 1987. China's nomenklatura system. *Probs. Communism*, 36, p.36.

Caselli, Francesco, and Massimo Morelli, "Bad Politicians," *Journal of Public Economics*, 88 (2004), 759–782.

Chen, Ye, Hongbin Li, and Li-An Zhou. 2005. "Relative Performance Evaluation and the Turnover of Provincial Leaders in China." *Economic Letters* 88 (3): 421–25.

Chen, Ting, and James Kai-sing Kung. 2016. "Do Land Revenue Windfalls Create a Political Resource Curse? Evidence from China." *Journal of Development Economics* 123: 86–106.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. "Who becomes a politician?." *The Quarterly Journal of Economics* 132, no. 4 (2017): 1877-1914.

Dixit, Avinash, Gene M. Grossman, and Faruk Gul. "The dynamics of political compromise." *Journal of Political Economy* 108.3 (2000): 531-568.

Edin, Maria. 2003. "State Capacity and Local Agent Control in China: CCP Cadre Management from a Township Perspective." *China Quarterly* 173: 35–52.

Egorov, Georgy, and Konstantin Sonin. "Dictators and their viziers: Endogenizing the loyalty–competence trade-off." Journal of the European Economic Association 9, no. 5 (2011): 903-930.

Frye, Timothy, and Andrei Shleifer. 1997. "The Invisible Hand and the Grabbing Hand." *American Economic Review* 87 (2): 354–58.

Gandhi, Jennifer, and Adam Przeworski. 2006. "Cooperation, Coop- tation, and Rebellion under Dictatorships." *Economics & Politics* 18 (1): 1–26.

Grover, A. and Leskovec, J.,"node2vec: Scalable feature learning for networks," in Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. *ACM*, 2016, pp. 855–864.

Guo, Gang. 2007. "Retrospective Economic Accountability under Authoritarianism: Evidence from China." *Political Research Quarterly* 60: 378–90.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13(4): 631–68.

Jia, Ruixue, and Yiqing Xu. 2018. Rotating to the Top: How Career Tracks Matter in the Chinese Communist Party. Working Paper.

Jiang, Junyan and Zeng, Yu, Countering Capture: Elite Networks and Government Responsiveness in China's Land Market Reform (December 13, 2018). Forthcoming, *Journal of Politics.*

Jiang, J., 2018. Making bureaucracy work: Patronage networks, performance incentives, and economic development in china. *American Journal of Political Science*, 62(4), pp.982-999.

Jin, H., Qian, Y. and Weingast, B.R., 2005. Regional decentralization and fiscal incentives: Federalism, Chinese style. Journal of public economics, 89(9-10), pp.1719-1742.

Keller, F. B. (2016). Moving beyond factions: Using social network analysis to uncover patronage networks among Chinese elite. Journal of East Asian Studies, 16, 17-41.

Landry, Pierre F. 2008. Decentralized Authoritarianism in China: The Communist Party's Control of Local Elites in the Post-Mao Era. Cambridge: Cambridge University Press.

Landry, P.F., Lü, X. and Duan, H., 2018. Does performance matter? Evaluating political selection along the Chinese administrative ladder. Comparative Political Studies, 51(8), pp.1074-1105.

Li, C. (2001). China's leaders: the new generation. Rowman & Littlefield.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Con- trol in China." Journal of Public Economics 89 (9/10): 1743–62.

Lu, Qinglian. 2018. Bureaucracy and Networks : the Politics of Career Mobility in Organizational Systems. Ph.D. Thesis, Stanford University.

Lü, X. and Landry, P.F., 2014. Show me the money: Interjurisdiction political competition and fiscal extraction in China. American Political Science Review, 108(3), pp.706-722.

Manion, M., 1985. The cadre management system, post-Mao: The appointment, promotion, transfer and removal of party and state leaders. The China Quarterly, 102, pp.203-233.

Maskin, Eric, Yingyi Qian, and Chenggang Xu. 2000. "Incentives, Information, and Organizational Form." Review of Economic Studies 67 (2): 359–78.

Meyer, D., Shih, V. C., & Lee, J. (2016). Factions of different stripes: Gauging the recruitment logics of faction in the reform. Journal of East Asian Studies, 16, 43-60.

Miller, A. L. (2004). Hu Jintao and the party politburo. China Leadership Monitor, 9:1–12.

Oi, Jean C. "Communism and clientelism: rural politics in China." World Politics 37, no. 2 (1985): 238-266.

Oi, J. C. (1999). Rural China takes off: Institutional foundations of economic reform. Berkeley: University of California Press.

Oksenberg, Michel. "The exit pattern from Chinese politics and its implications." The China Quarterly 67 (1976): 501-518.

Opper, Sonja, Victor Nee, and Stefan Brehm. 2015. "Homophily in the Career Mobility of China's Political Elite." Social Sci- ence Research 54: 332–352.

Qian, Y. and Weingast, B.R., 1997. Federalism as a commitment to reserving market incentives. Journal of Economic perspectives, 11(4), pp.83-92.

Razo, A. (2008). Social foundations of limited dictatorship: networks and private protection during Mexico's early industrialization. Stanford University Press.

Scott, J. C. 1972. Patron-client Politics and Political Change in Southeast Asia. The American Political Science Review, pages 91–113.

Stokes, Susan C. Perverse accountability: A formal model of machine politics with evidence from Argentina. American Political Science Review, 99(3):315– 325, 2005

Svolik, Milan W. "Power sharing and leadership dynamics in authoritarian regimes." Amer- ican Journal of Political Science 53, no. 2 (2009): 477-494.

Walder, Andrew G. "Career mobility and the communist political order." American sociological review (1995): 309-328.

Weber, Max. [1921] 1958. “Characteristics of Bureaucracy.” In *From Max Weber: Essays in Sociology*, ed. H. H. Gerth, and C. W. Mills. New York: Oxford University Press.

Whiting, Susan H. 2004. “The Cadre Evaluation System at the Grass Roots: The Paradox of Party Rule.” In Holding China Together: Diversity and National Integration in the post-Deng Era, ed. D. L. Yang, and B. Naughton. New York: Cambridge University Press, 101–19.

Yao, Yang, and Muyang Zhang. 2015. “Subnational Leaders and Economic Growth: Evidence from Chinese Cities.” Journal of Economic Growth 20(4): 405–36.

Zhao, Wei, and Xueguang Zhou. "Chinese organizations in transition: Changing promotion patterns in the reform era." Organization Science 15, no. 2 (2004): 186-199.

Zhou, Xueguang. (2001). Political Dynamics and Bureaucratic Career Patterns in the People’s Republic of China, 1949-1994. Comparative Political Studies, 34(9), 1036–1062.

Zhou, Xueguang. (2016). The separation of officials from local staff: The logic of the Empire and personnel management in the Chinese bureaucracy. Chinese Journal of Sociology, 2(2), 259–299.

周雪光, 艾云, 葛建华, 顾慧君, 李兰, 卢清莲, 赵伟 and 朱灵, 2018. 中国地方政府官员的空间流动: 层级分流模式与经验证据. *社会*, 38(3), pp.1-45.

Appendices

(R and Python Code to Clean Data and Reproduce Main Results)
Starting from Next Page