The Innovation Cost of Short Political Horizons: Evidence from Local Leaders' Promotion in China*

Xueping Sun †

Stanford Graduate School of Business

This version: November 19, 2023
[Click here for the latest update].

Abstract

This paper examines how politicians' time horizons affect the choices between policies that yield short-versus long-term growth. I digitize the career histories of Chinese city leaders, link them to economic policies and innovation outcomes, and exploit political connections formed through previous work ties to generate variation in leaders' promotion expectations. I find that when leaders are connected, they can rationally expect an earlier promotion. Such expectations lead them to pursue a fast-over-slow strategy for growth: higher spending on infrastructure, lower spending on science and technology, and lower effort in promoting innovation. As a result, the local economy has higher short-term growth but lower future patenting and long-term growth.

Keywords: Political Horizon, Innovation, Tenure Length, China Political Economy

*Acknowledgments: I thank Philippe Aghion, Ulrika Ahrsjö, Ingvild Almås, Guillem Bagaria Portet, Tessa Bold, Konrad Burchardi, Mark Cohen, Divya Dev, Mitch Downey, Jonathan de Quidt, Ruixue Jia, Roza Khoban, John Kramer, Per Krusell, Kieran Larkin, Evelina Linnros, Jens Oehlen, Laia Navarro-Sola, Arash Nekoei, Chloe Nibourel, Peter Nilsson, Elle Parslow, Torsten Persson, David Schönholzer, Victor Shih, Josef Sigurdsson, Fabian Sinn, David Strömberg, Gustavo de Souza, Ferenc Szucs, Anna Tompsett, Ekaterina Zhuravskaya and seminar participants at the IIES, Stockholm University, 21st Century China Center at UCSD, Gothenburg University, SIOE 2022 and EEA 2022 for helpful comments and suggestions.

†xpsun@stanford.edu

1 Introduction

Promoting innovation-driven growth has become the central economic policy issue for many governments in the world (Aghion and Howitt, 2006; Green et al., 2010; Bloom et al., 2019). But an often neglected aspect among existing discussions is the tension between innovation-driven growth and politicians' horizons: innovation is slow to pay off while politicians come and go at a relatively high frequency. Such tension might lead to public under-investment in innovation and compromise long-term growth.

This study examines the effect of city leaders' horizon on innovation in China, where the tension is particularly relevant amid China's national goal of "innovated in China" after four decades of growth (Naughton, 2017; Wei et al., 2017). The local leaders, in particular at the city level, are key to carrying out this national agenda. They run the bulk of the economy and have enormous economic resources at their disposal to advance policies with discretion (Xu, 2011).[1] However, their incentives might be governed by short-term pursuits. Overall, their tenures are short and unpredictable — since 1992, the average tenure of city leaders has been about 3.5 years. Upon the end of the term, city leaders are usually assigned to another location and moved either one level up or sideways in China's leadership hierarchy. Given that economic performance during their terms is closely associated with promotion (Li and Zhou, 2005; Xu, 2011; Jia et al., 2015), city leaders might pursue short-term growth at the cost of innovation and long-term growth.

I build a novel dataset that links city leaders' careers to economic policies and innovation outcomes, covering all of China's 283 prefecture-level cities from 2000 to 2018.[2] First, I digitize the universe of local leaders' CV profiles from multiple sources, which include 3214 city leaders and 318 provincial leaders, and develop a novel method to structuralize the raw CVs using natural language processing. Next, I combine both local governments' fiscal spending and annual work report using text analysis to measure the policy trade-off of fast-versus-slow, by comparing infrastructure development with innovation investment. Third, I compile the universe of Chinese patents of invention using data from the European Patent Office and the China National Intellectual Property Administration and use patenting activity at the city-year level to measure innovation.[3]

[1] The local governments spend 87% of the total government budget and own 47% of state-owned-enterprise assets. They can also finance economic development through off-balance-sheet borrowing by using local government financing vehicles (LFVs).

[2] There are 333 prefecture units: 17 prefectures, 30 autonomous prefectures (ethnic minority regions), 3 leagues (in Inner Mongolia) and 283 prefecture-level cities. The data on local government spending and local government work reports are available only for the prefecture-level cities.

[3] The patent of invention is also known as the utility patent in America.

I exploit political connections to generate shocks in city leaders' tenure expectations. Previous studies have shown that political connections can shorten tenure (Kou and Tsai, 2014; Pang et al., 2018), enhance promotion (Jia et al., 2015; Shih et al., 2012) and elicit effort in the Chinese political leadership hierarchy (Jiang, 2018). Building on these studies, I define political connection as an indicator of pre-determined social ties between city leaders and provincial incumbents. Thus, the variation in connection is driven by the turnovers of provincial leaders and the pre-existing ties between a city leader and the provincial incumbent. While the former is likely to be exogenous, as the national leaders control the turnover of provincial leaders, the latter might be prone to endogenous selection, if those with connections are more likely to become city leaders in the first place. To address this concern, I separate the treatment effect from the selection effect of connections by controlling for whether city leaders start as connected upon appointment.

To measure political connections, I leverage leaders' career networks to construct pre-determined proxies of connectedness and focus on a specific type of work ties: previously formed direct subordinate-superior relationships, i.e., whether the current superior was also a direct superior during any previous work spells. This measurement not only allows for a more accurate depiction of patronage relationships at a greater granularity but also avoids sparsity as subordinate-superior ties are more common than other types of social ties. To implement this measurement, I develop a fast and scalable search method using matrix representation.

My empirical analysis shows that connected leaders can rationally expect earlier promotion. In line with the descriptive work by Kou and Tsai (2014), I find that a connected leader's tenure is 12% shorter than that of an unconnected leader, which is equivalent to 4 months. In addition, their promotion likelihood is 20% higher, of which the magnitude is comparable to previous studies on connections enhancing promotions (Shih et al., 2012; Jia et al., 2015). I also find that connections help old and young leaders differently: while being connected helps old leaders attain promotion more quickly, it only shortens young leaders' tenure without increasing the promotion likelihood. I leverage this differential effect to isolate the horizon effect from the promotion effect in my analysis.

I observe a pursuit of fast policy over slow policy among connected leaders. On the one hand, their government spending on infrastructure development is 6.8% higher, amounting to an annual increase of 26.8 million RMB. On the other hand, their government spending on science and technology is 10% lower, equivalent to a reduction of 10.4 million RMB. Moreover, cities with connected leaders put roughly 0.07 standard deviation (SD) less emphasis on innovation, which suggests that a lower priority is given to promoting innovation.

Furthermore, this fast-over-slow trade-off is primarily driven by young leaders.

In line with the findings above, I also observe that higher short-term growth is achieved at the cost of compromising long-term growth. The immediate effect of connections on gross domestic product (GDP) growth is 1 percentage point (pp), which is 8% relative to the average growth rate. However, after the effect on GDP growth peaks in the second year, it decays and disappears by the fifth year, and becomes negative (-1 pp) between the seventh year and the ninth year. Meanwhile, I observe a persistent decline in patenting activity in the future that lasts beyond the seventh year. The immediate effect of connection on patent counts is a 12% reduction. Even after the seventh year there still exists an 8% gap. I argue that the sizable impact on innovation and growth is likely to be a combination of direct change in the policy environment and indirect change in the market's environment. Consistent with this argument, I find that the decline occurs for both universities and firms, among which the former rely crucially on direct funding support from the government and the latter's innovation activities are more responsive to market dynamics.

While these results are consistent with the hypothesis of short-horizon, these findings also have alternative interpretations. First, political connections might affect policies through other channels as well, for example, through resource transfer. I examine both the explicit fiscal transfers and implicit transfers such as credit expansion and do not find differences between connected leaders and unconnected leaders. Second, infrastructure development differs from investment in science and technology in many aspects other than the former being faster to pay off than the latter and the observed trade-off can be driven by other factors such as risk-taking and rent-seeking. While my analysis cannot fully reject these alternative hypotheses, I do provide evidence that the pursuit of fast-over-slow driven by a short horizon is the most likely channel. The four sets of findings in this research — namely connected leaders stay shorter, they pursue infrastructure over innovation, their jurisdictions have higher growth in the short run but lower growth in the long run, and all these patterns are primarily driven by leaders for whom only tenure is shortened by connections — cannot be entirely explained by rent-seeking or risk-taking. In addition, the findings of fast-over-slow are robust to alternative specifications and estimation strategies.

China's unique combination of economic decentralization and political centralization has sparked extensive discussions about the political economy of China's growth miracle and its future trajectory (Xu, 2011; Naughton, 2017). My paper contributes to this growing discussion by studying innovation policies and the prospective tensions between bureaucratic incentives and innovation-led growth. Current literature shares a wide consensus that local leaders' career concerns, created through political turnovers, are a double-edged sword in

China's rise as a world factory, contributing to both spectacular growth and serious economic problems (Maskin et al., 2000; Blanchard and Shleifer, 2001; Li and Zhou, 2005; Jia et al., 2015; Jia, 2017; Jia and Nie, 2017; Shi and Xi, 2018; Xiong, 2018). I add to this literature by showing that short tenures, as a consequence of frequent turnovers, lead to underinvestment in innovation and therefore might erect an institutional barrier to innovation-led growth.

Second, this paper complements existing studies on political short-termism by examining China and innovation. I show that political short-termism at the local level reduces innovation, despite the national leaders of China having long horizons. This discovery first contradicts the view that political stability enables long-termist policies in China by shedding light on the local level of political instability. It also adds to existing studies on the consequence of political short-termism in democracies on issues such as fiscal sustainability (Besley and Case, 1995), macroeconomic stability (Rogoff, 1990; Lindbeck, 1976; Alesina, 1988; Alesina and Perotti, 1996), legislative output (Dal Bó and Rossi, 2011; Titiunik, 2016), and corruption (Ferraz and Finan, 2011), by studying innovation. Moreover, while this paper reaches a finding that is similar to existing studies on innovation problems within firms, where short-termism widely exists (Porter, 1992; Lerner and Wulf, 2007; Aghion et al., 2013; González-Uribe and Groen-Xu, 2017), it also points to an unanswered but fundamental question for future research, which is how to make governments accountable for long-term policies such as innovation.

Finally, I also contribute a novel method to the future study of social connections. By developing a method of fast searching through CV networks, I am able to measure connections at greater granularity. This method differs from previous studies which have focused on social ties such as having attended the same universities (Xu, 2018; Shih and Lee, 2020), being from the same hometown (Xu, 2018; Li, 2019; Fisman et al., 2020; Shih and Lee, 2020), group affiliation (Fisman, 2001) or even ancestry (Xu, 2018). It also goes beyond previous studies on connection through work ties in the Chinese bureaucratic system that have focused on high-ranking politicians (Jia et al., 2015; Jia, 2017; Shih and Lee, 2020), or the most recently formed patron-client ties (Jiang, 2018) and leads to new findings on its role in advancing individuals' careers.

The rest of this paper is organized as follows. The next section briefly introduces relevant institutional features of local leaders' promotion and innovation policies in China. Section 3 describes the data collection and variable construction methods. Section 4 presents the empirical strategy and explains the econometric specification. Section 5 reports on the main results and Section 6 discusses the mechanisms and implications. Section 7 concludes.

2 Background

This section describes the institutional background of the political turnovers of local leaders and the innovation policies of local governments. For turnovers, I will focus on how the short tenure of local leaders is common and necessary, and how connections help the career progression of local leaders. For innovation policies, I will focus on the capacity and discretion of local governments.

2.1 Political Turnovers and the Role of Political Connections in China

To move up the political career ladder, fast turnover is necessary as many steps have to be taken to reach the top before hitting the age limit. Local Chinese leaders are political appointees and bureaucrats: they are appointed by their superiors, but their careers are governed by bureaucratic rules. These leaders progress through their careers by moving up the hierarchy ladder; decisions regarding their progression are made by their direct superiors, who are leaders at the level immediately above the local leaders. The progression has to be made step-by-step, and many steps have to be taken to reach the top. For example, to rise from an entry-level bureaucrat position to that of a city leader, there are six steps, and it takes an average of 30 years to reach that position. An additional four steps are required for a city leader to become a national leader. If a local leader has not reached the next level before she hits the corresponding age limit, then she retires. This practice dates back to the late 1980s and was formally institutionalized in the mid-1990s.[4][5] As a result of these rules, short tenures are necessary for career success, and fast promotions are widely observed (Kou and Tsai, 2014).[6] For example, all current politburo members, who run the highest decision-making organ of China, were on a fast track: along each step of the bureaucratic hierarchy, they arrived there at a younger age, stayed at each position for a shorter period of time, and were promoted earlier.[7]

Both connections and economic performance help the career progression of local leaders.

[4] *Notice about Revising the List of Cadres Positions Managed by the the Central Committee of the Chinese Communist Party (CCCPC).*, issued by the organization department of the CCCPC in July 1984.

[5] *The Interim Regulations on Selecting and Appointing Leaders of Party and Government.* issued by the CCCPC in February 1995, *The Regulations on Selecting and Appointing Leaders of Party and Government.*, issued by the CCCPC in July 2002, and *The Regulations on Selecting and Appointing Leaders of Party and Government (Revised Version).*, issued by the CCCPC in 2013 and 2019. There are extensive discussions on the causes and implications of this unique combination of political appointees and the bureaucratic career path. In this article, I will focus on the time dimension of leaders' career paths.

[6] Kou and Tsai (2014) interpret fast promotion as a choice made to balance stability and rejuvenation.

[7] For instance, the current president of China, Xi Jinping, became the party secretary of Fuzhou, the capital city of Fujian province, at a recorded young age of 37.

The CCCPC's guideline is "equal importance of competence and political integrity".[8] In practice, both economic performance and political connection are found to be associated with promotion (Li and Zhou, 2005; Shih et al., 2012; Jia et al., 2015; Wong and Zeng, 2018). Moreover, connections play a crucial role in putting leaders on the fast track. Pang et al. (2018) find a dual-track system of cadre management: some cadres are preferred, cultivated, and promoted along a fast track, while others have to earn their promotions alone the regular track.[9]

2.2 Capacity, Policy and Innovation

Innovation policies in China, like any other kind of economic policy, follow a top-down approach in which the governments at the upper-level set the goals and guidelines and the governments at the lower level implement them.[10] Local governments have considerable economic resources at their disposal to advance such policies. For example, the subnational government's share of government spending in 2016 was 85.4%, of which the prefecture-level government's share was 58.2%.[11] For innovation, the subnational government share of government R&D spending was approximately 60%. The subnational government also controls a major share of the economic resources. In 2016, the subnational government's asset of the state-owned-enterprise (SOE) was 59.5 trillion Yuan, which amounts to 47% of all SOE assets.[12] Using these resources, local governments can promote innovation in three ways. First, they can coordinate the development of R&D projects. Examples include creating platforms for firms to collaborate, or building science parks for incubating high-tech start-ups. Next, they can finance innovation by offering fiscal subsidies and financial investment funds. Last but not the least, they can also implement complementary measures such as tax refund, credit support, and preferred public procurements.

The decentralized implementation of innovation policies leaves local governments with considerable discretionary powers. The local government can, for example, decide which firms should enter science parks and receive tax refunds or other kinds of policy support,

[8] *The Interim Regulations on Selecting and Appointing Leaders of Party and Government*

[9] They argue that the Chinese Communist Party (CCP) is strongly ambivalent toward institutionalization: the CCP is fond of using institutional rules but is also afraid of being trapped by them. As a result, the regime applies its rules selectively to some cadres but not to others, leading to "sprinting with small steps" and thus a dual-track system of cadre promotion. The dual-track design illustrates how the regime employs the regular track to mobilize efforts and the fast track to cement loyalty.

[10] For a detailed review, see Ding and Li (2015). China started promoting innovation systematically in 1999 and launched the Medium-and-Long-term National Plan for Science and Technology Development (2006-2020) in 2020.

[11] http://www.gov.cn/xinwen/2017-03/06/content_5173807.htm.

[12] With financial firms excluded, see http://www.gov.cn/xinwen/2016-09/13/content_5107843.htm.

and which universities should get funding for research. They can also decide to offer none of these forms of support. I now illustrate how local governments are involved in innovation using the example of the city of Dongguan.[13] Between 2006 and 2010, Dongguan initiated the *Technology Dongguan* program and spent 5 billion Yuan, which is 1.8% of its fiscal budget, to promote innovation. This program subsidizes firms' technology upgrades, the credit supply to small and medium-sized enterprises, projects granted by upper-level government funds, and innovation services. Besides directly financing innovation, Dongguan also built the Songshan Lake Science (DGSSL) and Technology Industrial Park in what used to be hard-to-reach rural areas.[14] The construction of DGSSL involved a massive infrastructure investment on public facilities, including a new highway linking it to the main traffic artery of Dongguan. On top of public facilities, DGSSL prioritizes entrance for high-tech firms with a set of supportive policies. Formally opened in 2010, by 2019 this technology park had attracted more than 800 IT firms, 400 firms that manufacture intelligent robots, and 300 biotech firms, and had filed 4869 patent applications.[15] However, not all science parks in China are as successful as DGSSL. For example, researchers find aggregate level evidence on the misallocation of R&D funds in China (König et al., 2020).

Last but not the least, the government's innovation policies matter to a different extent for innovation by firms compared to that by universities. While firms' R&D activities are mostly self-financed, their innovation activities are dependent on continued support to the government policies. For example, a survey conducted by the National Bureau of Statistics (NBS) on firm innovation activities during the period 2002-2006 reports that close to 50% of the firms with innovation activities find the following policies to be helpful: tax discounts, credit support, government procurements, intellectual property protection, industrial policies, trade policies, and talent recruitment.[16] In the most recent survey, more than 70% of the firms confirm the usefulness of these policy interventions.[17] Universities and government-affiliated research institutes, in contrast, rely heavily on direct government funding. For example, a survey on universities' innovation activities in 2014 reports that universities receive 60% of their R&D funding from the government and spend around 30% of their R&D funding on basic research and 50% on applied research.[18]

In short, the incentives of innovation policies can vary significantly across regions and

[13] Dongguan is famous for its nickname of the "world factory" and as the R&D headquarters of the company Huawei.

[14] http://ssl.dg.gov.cn.

[15] http://ssl.dg.gov.cn/zjyq/yqgk/cyfz/index.html.

[16] http://www.stats.gov.cn/ztjc/ztsj/2006gysj/200802/t20080222_61480.html.

[17] http://www.stats.gov.cn/tjzs/tjsj/tjcb/dysj/201704/t20170421_1487052.html.

[18] Tables 4-2 from http://www.most.gov.cn/zxgz/cxdc/cxdcjcbg/201710/P020171027326121716670.pdf.

sectors, and the factors driving these differences are unclear. The reason behind the varying levels of innovation success in different parts of China is precisely what this study attempts to uncover.

3 Data and Measurements

I study innovation for all of China's 283 prefecture-level cities from 2000 to 2018. This sample is primarily chosen due to its relevance and data availability. I focus on the city-level governments as they are significant players in economic policies, both in terms of responsibilities and discretionary powers over economic resources (Xu, 2011). I restrict my analysis to the period 2000-2018 due to data availability constraints. For the period before 2000, I lack leader features for about 15% of the cities. From 2018 onwards, there is data truncation due to the time lag in reporting patent applications.

My dataset consists of leaders' CV profiles, local government spending and GWRs, and invention patent applications. Aside from these four main components, I collect data on city yearbook statistics for local economy outcomes; the variables include GDP, investment, industry revenue, population, and others. I now describe the collection of CVs, government spending and GWRs, and patent data in detail.

3.1 Data Collection

3.1.1 Leaders' CVs

First, I construct a database of the Chinese local leaders' demographic characteristics and career trajectories using their CVs. There are two public sources of local leaders' career backgrounds in China: the local leader database from China Stock Market & Accounting Research (CSMAR) and the Chinese Local Government Officials Database (CGOD) from the Chinese Research Data Services (CNRDS). However, even a combination of these two sources is incomplete, and it is common to miss data on leaders' early careers, which is particularly problematic for this study. To approximate the universe of the Chinese local leaders' CVs as closely as possible, a significant amount of effort is spent on collecting and validating data.

Then, I use the CV data to construct a panel data set that contains the time series of the positions of each leader. For this step, I characterize each position by its location, office and title, all parsed using the NLP-NER method. For example, for the deputy mayor of

Wuhan, the office is "government", the title is "vice head", and the location is "Wuhan". So far, my database contains complete CV information for 3443 leaders, which includes 318 provincial leaders and 3,214 city leaders.[19] This step is crucial for measuring a leader's political connections, as explained later on. For all prefectures in the period 2000-2018, the universe of city leaders contains 3,325 individuals and I have CV characteristics for 92.15% of these individuals.

Using the structured CV data, I also measure aspects other than the political connections, such as career background and promotion outcome. For career background, the variables are age, education background in STEM, work experience in STEM, seniority (accumulated working time in government sectors), and progression along the fast track by being promoted within every 3.5 year period. I define the promotion outcome as a dummy variable capturing if a leader's next position in the subsequent 3 years has either a higher nominal rank or the same nominal rank but a higher de-facto rank and is not a "retirement" position.[20] Positions with "a higher nominal rank" than that of city-level leaders are sub-provincial level positions. Examples include standing committee members of provincial-level party committees or the mayors of sub-provincial cities. I also consider a move to a position with "the same nominal rank but a higher de-facto rank" as a promotion.[21] "Retirement position" refers to a position at the People's Congress (PC) or the Committee of People's Political Consultative Conference (PPCC) at the city level or the province level, as positions in these two organizations are likely indicators of (early) retirement.

3.1.2 Local government spending and work reports

It is impossible to fully account for local governments' various economic policies that affect innovation directly or indirectly. I instead focus on policy measurements that reflect local governments' engagement in promoting innovation by combining both government spending data and policy texts. Specifically, I focus on the infrastructure development policies that immediately stimulate growth, and on science and technology, that affect growth slowly.

I collect government spending data from two sources. For spending on science and technology, I use the China Premium Database provided by the CEIC. The sample covers all

[19] Eighty-nine of them became provincial leaders later on.

[20] It is common that at the end of the current term, a city leader first gets a flat move first but shortly afterwards obtains a one-level-up move. Using a forward window of 3 years can rule out such transitory moves.

[21] The set of such positions includes heads of departments within the provincial party organization or the provincial government who are responsible for personnel decisions and economic decisions. A more detailed description is provided in the appendix.

Chinese prefecture-level cities from 2003 onwards.[22] For spending on infrastructure, I use the China Urban-Rural Development Database provided by EPS Data, for which the source data are from MHUD. The sample period for Chinese prefecture-level cities is 2006-2016.

I also collect local government work reports (GWR) to account for complementary policies on innovation. I collect 4100 GWRs from city governments during the period of 2006-2018 from People Data, a database owned by the newspaper *People's Daily* in China.[23] I then measure local governments' policy engagement in promoting innovation at the city-year level using text analysis.

3.1.3 Invention patent application data

I combine patent application data from the Patstat database of the European Patent Office (EPO) and the PatViewer database, which is indirectly owned by the China National Intellectual Property Administration (CNIPA). Both databases have used data from CNIPA since 1985. From Patstat, I collect information on the patent attributes, such as authorization, International Patent Classification (IPC) code, citations, and technical fields. From PatViewer, I collect data on patent applicants' attributes, such as name and address. I focus on invention patents by Chinese applicants. There are a total of 10.11 million CNIPA invention patent applications documented by Patstat. Among these, I identify 8.73 million applications with a Chinese address.

3.2 Measuring Connections Using Previous Subordinate-superior Ties in CVs

I create a dummy indicator of the existence of any previous subordinate-superior ties between a city leader and the provincial leaders to measure connections. "Previous" refers to the past career paths taken before an individual becomes a city leader. "Subordinate-superior" refers to whether the person as a subordinate work with another person as the superior. In such ties, the superior, as the one-level-up leader, is responsible for the appointment, evaluation and promotion of the subordinate.

More specifically, I define a subordinate-superior tie as a pair of individuals simultaneously positioned at two vertically adjacent nodes in the Chinese bureaucratic hierarchy, with the person at the higher-ranked node arriving before the person at the lower-ranked node.

[22] These data have been available from 2003 onwards, after the budget management reform. Before 2003 the spending on science and technology was contained in the category of government spending on science, education, culture, and health.

[23] Before 2006, digital versions of work reports were not available for the majority of the local governments.

Due to the M-shape of the Chinese bureaucratic system (Qian et al., 2006), there exist three types of subordinate-superior ties. The first is within-location-within-organization, such as the head of the anti-corruption bureau and the division head within the anti-corruption bureau in the same region. The second is within-region-between-organization, for which two scenarios are possible. In the first scenario, a department-affiliated organization is supervised by the department to which it is affiliated. For example, the procuratorate in the same region oversees the anti-corruption bureau. In the second scenario, a department within the local government/party is supervised by the local government heads (vice-heads) and the local party committee. For example, the head of government (mayor) is the direct superior of all of the department heads in the local government. The third is between-region-within-organization. For example, the anti-corruption bureau of the province of Hunan oversees the anti-corruption bureau of Changsha, which is its provincial capital city.[24]

This definition of the subordinate-superior tie is highly restrictive in two aspects. First of all, it only accounts for a relationship between two individuals without any intermediaries, by restricting it to "vertically adjacent nodes simultaneously". Second, it only accounts for a tie formed by the superior appointing the subordinate, and not by the superior inheriting the subordinate, by restricting it to the superior arriving before the subordinate. These two restrictions allow for an accurate depiction of a direct patronage relationship.

The implementation of this measurement involves three steps.[25] The first step is to characterize each CV as a sequence of job events using the NLP-NER method. Each job is characterized as a combination of location (where the job is), organization (which section of the bureaucratic system), and title (the level of the position). This characterization is crucial for identifying the hierarchical relationship between jobs. The second step is to create a matrix of position hierarchies based on administrative rules. With these two steps of preparation, I can then search for subordinate-superior ties based on whether two individuals work in the same organization, in which one person holding a job oversees that of another person. I implement this step by developing a speedy search method using matrix representation. By construction, the on-and-off status of a leader's connection status is driven by the turnovers of provincial leaders. Thus, when the central party committee changes provincial leaders, a city leader's connection status will either stay the same, or switch off or on, depending on whether previously formed subordinate-superior ties exist between her and the provincial leaders.

[24] Since 1984, the management of cadres has followed two rules: appointments are made by the party committee one level up, and they are subject to the leadership both of the local authority (in the local government or the local party committee) and of the same department at the upper level.

[25] Appendices A and B show the details.

3.3 Measuring Policy Posture Using Local GWRs

I use each local government's annual government work report (GWR) to measure its policy posture on innovation-related and other issues. I define the policy posture as the share of content dedicated to a specific topic in each report. I measure content at the sentence level and compute the share of a topic using the ratio of the total length of sentences belonging to a specific topic over the total length of the policy document.

This differs from the bag of words (BOW) approach used by existing studies. From the rudimentary word counts to more sophisticated methods such as topic modeling, there have been substantial developments in the study of local governments' policy agendas using GWR (Jiang, 2018; Campante et al., 2019). One thing that these studies have in common is to treat each document as a BOW. The difference, depending on the exact method used, arises from how to count the relevant words. The BOW approach is appropriate for documents with a single theme, such as user reviews, US congressmen's speeches and news articles (Gentzkow et al., 2019). The Chinese GWRs fall outside this scope, as they have a highly homogeneous structure and have multiple themes, which results in either low cross-sectional variation or an *ad-hoc* interpretation of the topic model results. By exploiting the structural features of GWR, this study proposes a bag of sentences (BOS) model of GWR representation and uses sentences as the basic semantic unit.

I use text analysis to determine whether a sentence refers to innovation. Then, I aggregate the lengths of sentences on innovation. Finally, I compute policy posture by dividing the total lengths of sentences on innovation by the length of the document.

3.4 Measuring Innovation Using Invention Patent Applications

Innovation is intangible and hard to measure (Bloom et al., 2019). Patent count, however, is closely related to other measures of innovative activities (Shambaugh et al., 2017; Fang et al., 2020) and technology progress (Kelly et al., 2018), and is one of most commonly used proxies for innovation (Kogan et al., 2017).

I use invention patent applications instead of the granted invention patents for two reasons. The first is that this addresses the data truncation problem for recent patent applications as it usually takes 3-5 years for an invention patent to be granted. The second is that being granted is not the only measure for patent quality. To measure quality without truncating the data, I use the within-application-year citation ranking percentile to measure the quality of a patent. Specifically, for all invention patents applied from the same year from all around the world, I rank each patent based on the total number of citing simple DOCDB families it received.[26] A patent is labelled as high-impact if its within-application-year citation ranking is above the 90th percentile and low-impact otherwise.

I aggregate the patents at the city-year level to measure the quantity and quality of innovation. I first assign each applicant's address to the city in which the application was filed.[27] Then, I aggregate the total number of applications at the city level over alternative forward windows. The experiment of varying time windows is to account for the slow-to-yield nature of innovation activities (Aghion et al., 2018). In particular, I compare patents filed by firms to those filed by universities to study the heterogeneous effects.

3.5 Descriptive statistics

Table 1 summarizes the basic statistics of the key variables. The average tenure length of city leaders is less than 3.7 years and upon the end of their terms around one-third get promoted to a level up(Panel D). In terms of leaders' characteristics (Panel C), around 36% of them have a STEM background and 33% of them were on a fast track before becoming city leaders. The vast majority of them are at the age of 50-52 and therefore relatively safe from retiring directly. Connection with provincial leaders widely exists among city leaders: on average 68% of the time a city is under the governance of a connected leadership.

In terms of policies (Panel B), city governments spend nearly twice as much on infrastructure as on science and technology. The average annual budget spending on science and technology is around 533 million RMB or 3% of the total budget spending. Moreover, the average policy posture value on innovation is around 12.2, i.e. 12.2% of the GWR content is related to innovation. In comparison, local governments' average spending on infrastructure development is 1,076 million RMB, which is approximately 6% of the budget. These exact numbers are likely to underestimate the actual expenditure by local governments as they can use off-balance-sheet financial tools. However, as long as the actual spending is proportional to the on-budget spending, the issue of under-measuring should not affect the main conclusions of this study.

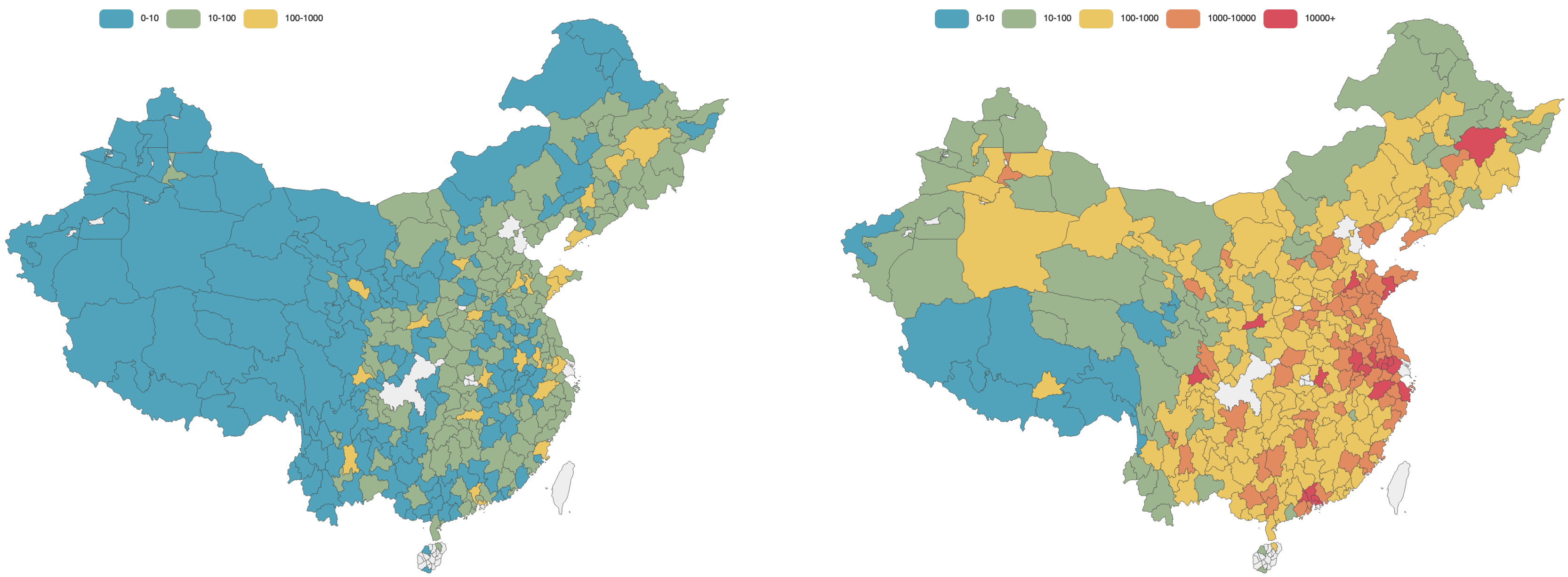
[26] The simple DOCDB family is a technical concept defined by the European Patent Office to represent patents that belong to the same set of technology range. For details, please go to https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families/docdb.html. The number of forward citing DOCDB simple families, representing the number of distinct future inventions citing the current invention, is widely used in existing literature to capture the quality of patents (Martinez, 2010)

[27] A small share of the observations has city names and boundaries that have changed over time. I assign these observations to the original cities.

Table 1: Summary Statistics

	stat					share of variation	
Variables	Count	Mean	Std	p5	p95	within	between
A: Innovation and Growth (city-year panel)							
Patent count	6015	989.80	3664.48	4.00	4453.00	0.56	0.44
Firm-patent count	6015	589.58	2664.50	0.00	2415.00	0.54	0.46
University-patent count	6015	154.14	701.09	0.00	719.00	0.54	0.46
GDP growth rate(%)	5171	13.09	8.85	0.60	26.55	0.94	0.06
B: Policy Outcome (city-year panel)							
Budget Spending on Infrastucture (million)	2575	1097.01	2320.50	50.49	4541.80	0.39	0.61
Budget Spending on Sci &Tech (million)	3658	533.27	1751.03	11.98	2064.90	0.47	0.53
Innovation Posture	3824	12.28	6.49	3.62	24.61	0.74	0.26
C: Leader Features (city-year panel)							
$Connected^{start}$	6209	0.80	0.40	0.00	1.00	0.78	0.22
$Connected^{startpsecretary}$	6090	0.62	0.49	0.00	1.00	0.76	0.24
$Connected^{startmayor}$	5935	0.57	0.50	0.00	1.00	0.75	0.25
$Connected$	6211	0.68	0.47	0.00	1.00	0.81	0.19
$Connected^{psecretary}$	6086	0.47	0.50	0.00	1.00	0.82	0.18
$Connected^{mayor}$	5959	0.50	0.50	0.00	1.00	0.80	0.20
$STEM^{psecretary}$	6269	0.37	0.48	0.00	1.00	0.77	0.23
$STEM^{mayor}$	6269	0.35	0.48	0.00	1.00	0.77	0.23
$FastTrack^{psecretary}$	6229	0.32	0.47	0.00	1.00	0.79	0.21
$FastTrack^{mayor}$	6133	0.29	0.46	0.00	1.00	0.79	0.21
$Age^{psecretary}$	6086	52.20	3.79	45.00	58.00	0.77	0.23
Age^{mayor}	5986	50.31	4.01	43.00	56.00	0.75	0.25
D: Turonver Outcome (finished city-leader spells)							
$TermLen^{psecretary}$	1935	3.69	1.77	1.08	6.92	0.81	0.19
$TermLen^{mayor}$	2078	3.42	1.66	1.08	6.25	0.74	0.26
$Promoted^{psecretary}$	1953	0.39	0.49	0.00	1.00	0.76	0.24
$Promoted^{mayor}$	1978	0.33	0.47	0.00	1.00	0.82	0.18

The descriptive statistics include the number of observations (count), mean, standard deviations (std), the 5 percentile, the 95 percentile, and the within/between city decomposition of variation. **Panel A** reports descriptive statistics for the outcomes of innovation and economic growth at the city-year level during 2000-2018. The patent count is the yearly number of patents filed, and firm-patent count and university-patent count are the yearly numbers of patents filed by firms and universities respectively, using data from EPO and CNIPA. GDP growth rate is the nominal growth rate measured in percentage points, using data from the city yearbooks. **Panel B** reports descriptive statistics for outcomes of policy choices at the city-year level. Budget spending on infrastructure is measured in million (RMB) and is available for the period of 2006-2016. Budget spending on science and technology is measured in million (RMB) and is available since 2003. Innovation posture is measured as the yearly percentage of contents dedicated to innovation in the GWR and this data is available for most cities since 2006. **Panel C** reports the summary statistics for leader features at the city-year level, including whether starting as connected ($Connected^{start}$), being currently connected ($Connected$), whether having a STEM background ($STEM$), whether was on a fast track ($FastTrack$) and age (Age). The superscript indicates whether the measurement is at the party secretary level ($psecretary$), at the mayor level ($mayor$), or at the leader team (the combination of party secretary and mayor) level in the absence of such superscript. **Panel D** reports the summary statistics on tenure length and promotion dummy among finished city-leader spells during the sample period. $TermLen$ is the length of tenure measured in years and $Promoted$ is a dummy and equals 1 if the leader gets promoted upon the end of the term. The superscripts are defined the same as those in panel C.



(a) # of application in 2000

(b) # of application in 2015

Figure 1: Patent applications of inventions across Chinese Cities: 2000 v.s. 2015

The data comes EPO and CNIPA. The empty areas on the map are the four provincial prefectures (Beijing, Tianjin, Shanghai and Chongqin), and several county-level cities which are outside the sample of prefecture-level cities.

For patenting activity, an average city in China files nearly 990 invention patents per year from 2000 to 2018, among which 590 patents come from firms and 154 comes from universities.[28] Patenting occurs widely across all Chinese cities but the distribution is highly skewed, as shown by figure 1. In 2000, there were only 24 cities (7.3% of the prefectures) that filed more than 100 applications. By 2015, there are 262 cities (78.4% of the prefectures) that file more than 100 applications, among which there are 97 cities that file more than 1000 applications. This pattern of widespread patenting is described as the "democratization of patenting" (Hu et al., 2017) and is closely associated with the surge of Chinese patenting at the aggregate level (Li, 2012; Wei et al., 2017; Hu et al., 2017). Meanwhile, the geographic distribution of patent applications is highly skewed, with patent applications by cities in the top 95th percentile being almost 1000 times that of cities in the bottom 5th percentile. Another feature of patent data is that the between-city variation explains around 40% of the variation, indicating a permanent gap in innovation across cities. To resolve the skewness of patent applications, I apply log transformation to all these outcome variables in regression analysis.

4 Empirical Design

This section discusses the identifying variation and the specification in detail. I aim at a causal estimate of leaders' tenure expectations on local innovation outcomes and use the political connection to generate exogenous shocks in leaders' tenure expectations. Furthermore, I separate the treatment effect from the selection effect of connections by controlling for whether city leaders start as connected upon appointment.

4.1 Identifying Variation

I use political connections to generate exogenous shocks in leaders' expectations of tenure length. The intuition is that when leaders are connected they might expect faster promotion, as existing studies have shown that political connections can put leaders on a fast track (Kou and Tsai, 2014; Pang et al., 2018) and improve their promotion prospects (Jia et al., 2015; Shih et al., 2012). A city leader is connected to her/his provincial leaders if they ever previously worked as a subordinate-superior pair (as defined in section 3.2). This implies that when a former patron enters into or exits from provincial leadership, a city leader's connection status switches on and off accordingly. The identification is thus driven by leaders

[28] The remaining category is individuals or individuals from small firms.

who change their connections during their respective terms. At the spell level, 33.6% of all leaders are never connected, 34.3% are always connected, and 32% are switchers.[29]

A natural concern is that leaders are selected into being connected. If innovation-prone leaders are more likely to be connected, the estimated effect of the horizon will be biased downward. Conversely, if innovation-averse candidates become city leaders because of connections, the estimated effect of the horizon will be biased upwards. Either way, leaders who begin as connected could systematically differ from those who do not. I examine the characteristics of leaders and cities based on whether the spell begins with a connection and find that leaders who begin as connected are similar to the remaining leaders, except that they have more upward connections (Appendix C.1) and are more likely to have worked as city leaders before.

I separate the treatment effect of connections from its selection effect by controlling for whether city leaders begin as connected.[30] As a candidate might become a city leader due to pre-existing work ties with the provincial incumbents in the first place, controlling for whether leaders are connected upon the appointment can thus hold constant the selection. Figure 2 illustrates the sample composition at the spell level based on how the connection switches within each spell. Depending on whether a leader starts as connected (C) or unconnected (N), and whether the leader is currently connected (C) or unconnected (N), a city-leader spell can experience one of the following four types of transitions: *NN*, *NC*, *CN* and *CC*. By holding constant whether a leader starts as connected, the average effect of connection then becomes a weighted average of the following two comparisons: *NN* versus *NC* (41% spells); *CC* versus *CN*(59% of the spells). In addition, this setting also extends Xu (2018) by being able to compare the treatment and selection effect of political connections.

Given the setup above, I make the following assumption for a causal estimate of the effect of connection: conditional on selection, the timing of switching connections is exogenous to outcomes. This assumption relies on the validity of the following two conditions. The first is that the initial selection is uncorrelated with potential outcomes. The second is that the

[29]If a leader ever gets a connection (66% of the spells), then it is most likely to happen at the beginning of the spell upon appointment (59% of the spells). Furthermore, of those who begin as connected, half of them will lose the connection over the years in term due to the turnovers of provincial leaders.

[30]An alternative is to control for leader fixed effects, which is not feasible in the setting of this study for two reasons. The first is that most leaders work as the mayor or the party secretary at the city level only once. In the sample period, only 7.2% of the mayors used to be mayors in another city, and 20% of the party secretaries used to be party secretaries in another city. Given that only around one-third of the spells are switchers, what remains is a small sample that is particularly problematic for studying the effect of connection on tenure length. The second reason is that each city is always co-governed by a mayor and a party secretary at the same time. Controlling for leader fixed effects also comes at the cost of removing the majority of the sample, which creates missing data problems when studying the dynamic effects.

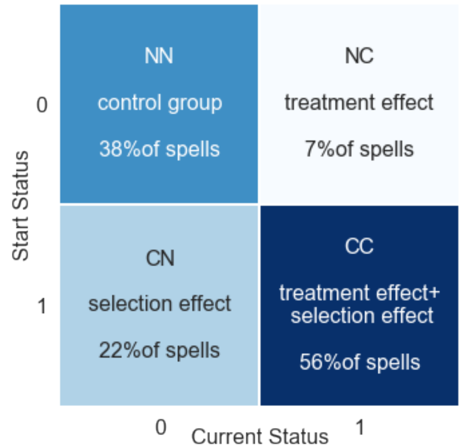


Figure 2: Sample composition by the switch of connections at the spell level

The unit of observation is the city-year spell for spells that start during the period of 1997-2017. The row represents the connection status at the beginning of the spell and the column represents the current connection status during the spell. Note that if a spell starts as unconnected, it can experience both an NC and an NN transition at different times within the spell. By analogy, for spells beginning as connected, the transitions CC and CN are not mutually exclusive. For spells that ever experience any CC transitions (56%), 34.3% of them experienced only a CC transition, i.e., always connected, and the remaining 22% experienced both a CC and a CN transition. The color intensity in each cell is proportional to the sample size of that cell.

switching of connections is uncorrelated with potential outcomes. I test the first condition by examining whether a city's initial economic conditions are correlated with its probability of being assigned to leaders beginning as connected. For the validity of the second condition, I examine whether a city's initial economic conditions are correlated with its probability of having leaders' connections status switched on or off after the appointment (Appendix C.2). In addition, I also test whether there exist any differential trends in economic growth and patenting for cities with connected leaders (Appendix C.3). Overall, I do not find any evidence that suggests that cities with connected leaders differ from those without in terms of innovation potential.

4.2 Reduced-form Specification

I examine whether connections affect the leaders' tenure and promotion expectations. Regressing turnover outcomes at the spell level on the current connection status face an issue of mechanical bias: the longer a leader stays, the more likely it is that the provincial leader changes and hence, the connection status switches. One solution is to estimate the effect of connection on tenure and promotion by comparing leaders with at least the same duration year by year.[31] The intuition is that for leaders who stay for at least the same duration, the likelihood of them experiencing a change in provincial leaders is comparable. In addition, this comparison allows me to directly link the results on leader turnovers to the results on city outcomes, as the latter are estimated at the city-year level. For tractability, I estimate a linear model of the hazard rate of leaders' tenure durations, specified as follows.

$$\begin{aligned}\pi_{i,c,t} =& \gamma_s * 1\{S_{i,c,t} = s\} \\ &+ \gamma * Connected_{i,c,t} + \gamma^{start} * Connected^{start}_{i,c} \\ &+ X_{i,c,t}\Gamma + \delta_c + \delta_t + u_{i,c,t} \end{aligned} \tag{1}$$

where i, c, and t are indexes for the leader, city and year, respectively. $\pi_{i,c,t}$ is an indicator of the turnover outcome. Regarding the outcome of exiting, $\pi_{i,c,t} = 1$ if the leader of the city c at year t ends his term.[32] Regarding the outcome of promotion, $\pi_{i,c,t} = 1$ if the leader of the city c at year t ends his term and is promoted. $S_{i,c,t}$ denotes the duration of the city leader since arrival in years, and $S_{i,c,t} = s$ if the leader has stayed for s years since being appointed. $X_{i,c,t}$ is a set of control variables that contains the initial value of the outcome variable and the local economic conditions before the leaders were appointed in terms of GDP, population, and the revenue of the manufacturing sector, and the indicators for the term year of the provincial leaders.[33] δ_c and δ_t capture city fixed effects (FE) and year FE, respectively.

The parameter of interest is γ. By construction, the coefficient γ measures the average difference in the probability of exiting (promotion) due to being connected, conditional on having stayed s years and after accounting for initial selection ($Connected^{start}$). The higher the probability of exiting, the shorter the tenure.[34]

[31] Another alternative is to estimate the effects of connection at the spell level by controlling for how long the provincial leaders last after the appointment, as shown in Appendix D.1.

[32] When a prefecture leader ends her/his term, she/he will either be reposted to a new position (but not necessarily promoted), or in rare cases, retire due to reaching the age limit.

[33] At the city level, a leadership team is newly appointed if either the mayor or the party secretary is replaced by a new one.

[34] The following formula computes the difference between always-connected and never-connected leaders in tenure:

$$\begin{aligned}\alpha^T &= E(T|Connected = 1) - E(T|Connected = 0) \\ &= \sum_{s=\tau_{min}}^{s=\tau_{max}} s(\gamma_s + \gamma) \prod_{s'|s' \leq s-1} (1 - \gamma_{s'} - \gamma) - \sum_{s=\tau_{min}}^{s=\tau_{max}} s\gamma_s \prod_{s'|s' \leq s-1} (1 - \gamma_{s'})\end{aligned}$$

I then investigate whether political connections affect a city's resource transfer, economic policies, growth, and innovation. The main equation takes the form:

$$\begin{aligned} y_{c,t} = \theta * Connected_{c,t} + \theta^{start} * Connected_{c,t}^{start} \\ + X_{c,t}\Theta + \delta_c + \tau_t + \epsilon_{c,t} \end{aligned} \tag{2}$$

where the subscripts c and t are indexes for city and year, respectively. $Connected_{c,t}$ represents connection at the team level, i.e., whether either the party secretary or the mayor is connected.[35] Similarly, $Connected_{c,t}^{start}$ denotes whether the leadership team begins as connected, i.e. at least one of two leaders begins as connected. The control variables $X_{c,t}$ contain the initial value of the outcome variable and the local economic conditions preceding the appointment of the leadership team in terms of GDP, population, the revenue of the manufacturing sector, and indicators for the term year of the provincial leaders.

Using the above specification, I explore the effects of connections on policy outcomes and innovation outcomes. I first examine the effects on economic policy and, in particular, quick-to-yield versus slow-to-yield policies. The outcome variables include annual growth in GDP, government spending on infrastructure, science & technology, and the local government's policy postures on innovation. I then investigate how connections affect growth and innovation outcomes dynamically. In all these regressions, the interest is θ, which measures the annual effect of connections.

5 Results

This section reports the baseline results on tenure, promotion, economic policy, growth, and innovation. Overall, connected leaders are promoted earlier, pursue fast investments over innovation policies, and bring about a persistent decline in future innovation.

5.1 Tenure, Promotion and Resource

I compare the average differences in tenure length and promotion rate that arise due to connections. To address the concern of the mechanical bias caused by the fact that lead-

[35] A city is always co-governed by a party secretary and a mayor, and it is unclear who is the de facto decision-maker. Thus, running the regression at the individual level by examining the impact of only one individual leads to omitting the other individual. It is possible to run regressions at the individual leader level by examining the effects of one leader's connection and the teammate's connections, but this generates the issue of self-reflection. For these two reasons, a reduced form estimated at the city level is the preferred specification.

ers who serve for longer are more likely to lose their connections, I use the year-by-year specification following equation (1). In this specification, *Connected* captures the average difference in exiting (promotion) due to being exposed to connection, by holding constant the selection effect ($Connected^{start}$). Table 2 presents the results and I use columns (4) and (7) for interpretation.

Table 2: Estimates of the effects of connections on tenure and promotion

	Exited				Promoted		
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Connected	0.0278***	0.0388***	0.0467***	0.0306***	0.0213***	0.0270***	0.0203**
	(0.009)	(0.011)	(0.011)	(0.011)	(0.007)	(0.008)	(0.008)
$Connected^{start}$	-0.0209***	-0.0342***	-0.0531***	-0.0449***	-0.0266***	-0.0350***	-0.0335***
	(0.008)	(0.010)	(0.011)	(0.011)	(0.007)	(0.008)	(0.008)
Old			0.0397***	0.0449***		-0.0377***	-0.0314***
			(0.007)	(0.007)		(0.005)	(0.006)
FastTrack			0.0398***	0.0423***		0.0233***	0.0267***
			(0.008)	(0.009)		(0.006)	(0.007)
Observations	14,647	14,720	14,720	12,139	14,712	14,712	12,134
R-squared		0.197	0.267	0.275	0.051	0.103	0.112
Mean	0.244	0.248	0.248	0.246	0.096	0.096	0.096
City and year FE			X	X		X	X
Controls				X			X
SE Cluster	City	City	City	City	City	City	City

The unit of observation is leader by year. The sample is a pooled sample of leaders, namely city mayors and party secretaries, who were appointed during the period 2000-2017 (2018 is skipped due to data truncation issues, as around 20% of the leaders appointed in 2018 still hold office). The dependent variable is an indicator of whether the leader leaves office (columns (1)-(4)) or is promoted (columns (5)-(7)) in a given year. The variable *Old* is a dummy and equals 1 if the leader's age upon appointment is higher than the median age of all leaders appointed in the same year. The variable *FastTrack* is a dummy that equals 1 if the leader's accumulated average tenure length across all her/his previous jobs is less than 3.5 years. All the regressions include a fully non-parametric baseline hazard for the number of years in office (job-tenure fixed effects (FE)). Column 1 shows results from a binomial regression with a complementary log-log model (marginal effects). The other columns show estimates using a linear probability model. Controls include the logs of GDP, population and industry revenue for the year before the local leader was appointed, and FEs for the term year of the provincial governor and the party secretary. Standard errors (SE) are clustered by city, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

First of all, I find that being connected significantly shortens leaders' tenure lengths. A connected leader has a significantly higher likelihood of exiting her current position at any time during the term (row (1), columns (1) to (4)). On average, a connected leader's exiting rate is 3.06% higher (row (1), column(4)), which is equivalent to a 12.4% higher likelihood than the average, or a reduction of 5 months (or 0.42 years). Moreover, this result is robust to alternative specifications. It is similar to the result obtained using binomial regression (column (1), whether controlling for city and year FE (column (2) v.s. column (3)) and control variables (column (3) v.s. column (4)).

Second, I observe that being connected significantly increases leaders' ex-ante promotion probabilities (row (1), columns (5) to column (7)). During any year of her/his term, a connected leader's promotion probability is 2% higher (row (1), column (7)), i.e., she/he has a 20.8% higher promotion likelihood relative to the average. This result is also robust to alternative specifications (columns (5)-(6)). However, as compared to Jia et al. (2015), the effect of political connection in this study has a smaller magnitude: Jia et al. (2015) find that connections, when measured as sharing the workplace with politburo members, increase the provincial leaders' promotion probabilities by 70%.[36] As I examine city leaders, the difference between this study and what was found by Jia et al. (2015) might be driven by the heterogeneous effect of connections by political rank.

Table 3: Estimates of the effects of connections on resources

	log(Total Fiscal Transfer)		Depdency on Debt Total Infra.Dev		Depdency on Pro.Gov Fiscal Infra.Dev	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Connected	0.0026	0.0193	0.0149	0.0109	0.0042	0.0075
	(0.016)	(0.019)	(0.009)	(0.010)	(0.007)	(0.007)
$Connected^{start}$	-0.0290	-0.0421	-0.0025	-0.0021	0.0065	0.0024
	(0.023)	(0.026)	(0.012)	(0.013)	(0.010)	(0.010)
Observations	5,239	4,826	4,948	4,453	1,521	1,394
R-squared	0.929	0.926	0.355	0.375	0.453	0.469
Mean	8.413	8.440	0.212	0.216	0.057	0.055
City and year FE	X	X	X	X	X	X
Controls		X		X		X
SE Cluster	City	City	City	City	City	City

The unit of observation is city by year. The dependent variable is the log values of fiscal transfer (columns (1)-(2)), debt dependency of infrastructure development (columns (3)-(4)), and fiscal dependendecy on provincial government for infrastructure development [a] (columns (6)-(7)). Due to data availability constraints, the sample period for fiscal transfer is 2000-2018, for the debt ratio of infrastructure development it is 2002-2018 and for fiscal dependency on provincial government for infrastructure development it is 2006-2016. Controls include the logs of GDP, population and industry revenue for the year before the local leader was appointed and fixed effects FE for the term year of the provincial governor and party secretary. Standard errors (SE), clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Infra. Dev.: infrastructure development; Pro. Gov.: provincial government.

[a] included both on-budget spending and off-budget spending.

Third, I find suggestive evidence of adverse selection due to connections (row (2)). Compared to leaders who are never connected, leaders who begin as connected but later lose their connections serve 6.8 months more (row (2), column (4)) and have a 34% lower likelihood

[36] Jia et al. (2015), Table 2, row (1), columns (1)-(3).

of being promoted (row (2), column (7)). One possible explanation is that these leaders are less competent but nonetheless become city leaders due to their connections. However, once the connections are lost, these leaders can no longer expect fast promotions and have to earn future promotions. Another explanation is that the lack of loyalty hinders their career progression, as connections to previous provincial leaders signal disloyalty. In either case, selection issues exist and should be controlled for when examining the impacts of connections on economic policies.

I interpret the effects of connections on turnover outcomes as the realization of the turnover expectation. However, connections can affect other aspects that change leaders' behaviors. For example, connected leaders might receive more support from provincial governments in explicit fiscal transfers or implicit transfers such as credit expansion. Table 3 presents the effects of connections on resources following equation 2. Overall, I find that the connections affect neither explicit nor implicit resource transfers. Cities with connected leaders do not receive more fiscal transfers (columns (1)-(2)), do not accrue more debt to finance local infrastructure development (columns (3)-(4)), and do not rely more on the provincial government's transfer for fiscal spending on infrastructure development (columns (5)-(6)).[37]

5.2 Economic Policies: Fast v.s Slow

Next, I examine how connections matter for economic policies that differ in the horizons required to yield policy impacts. Table 4 summarizes the results of the estimates on the impacts of political connections on the choice of fast v.s slow economic policies based on equation (2). In this specification, *Connected* captures the difference in policies when leaders are connected, conditional on initial selection ($Connected^{start}$). For the purpose of precision, I use columns (1), (3), and (5) for interpretation.

My analysis shows that connected leaders increase spending on infrastructure development (row (1), columns (1)-(2)). The local government's budget spending on infrastructure is 6.87% higher, which is equivalent to a 26.8 million RMB increase.[38] Although this effect is not very precisely estimated (row (1), column (1)), it is robust to adding initial values and the initial trend of budget spending on infrastructure before leaders' arrival (row (1),

[37] The city government can issue off-balance-sheet debt through the local financing vehicle (LFV) companies. LFVs are owned by local governments and borrow and spend on behalf of local governments. For a detailed institutional context on the LFVs and how the local governments use them for off-balance-sheet financing, see Huang et al. (2020).

[38] The average annual budget spending on infrastructure by the city government is approximately 391 million RMB.

column (2)).

Table 4: Estimates of the effects of connections on economic policies

	log(Infrastructure)		log(Sci&Tech)		Innov.Posture (SD)	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Connected	0.0687*	0.0762	-0.0997***	-0.0923**	-0.0739**	-0.0717
	(0.041)	(0.060)	(0.032)	(0.046)	(0.037)	(0.052)
$Connected^{start}$	-0.0721	-0.1466*	0.0370	0.1005*	-0.0133	0.0175
	(0.049)	(0.082)	(0.041)	(0.060)	(0.049)	(0.070)
Observations	2,391	1,126	4,262	1,858	3,311	1,961
R-squared	0.865	0.895	0.935	0.918	0.707	0.683
Mean	6.103	6.402	4.581	5.713	-0.019	0.386
City and year FE	X	X	X	X	X	X
Controls	X	X	X	X	X	X
Init.Cond.Depvar		X		X		X
SE Cluster	City	City	City	City	City	City

The unit of observation is city by year and the sample period varies by outcome variable due to data availability constraints. The dependent variable is the log value of budget spending on infrastructure development (columns (1)-(2)), science and technology (columns (3)-(4)), and governments' policy posture on innovation (columns (5)-(6)). The sample period for budget spending on infrastructure development is 2006-2016, for budget spending on science and technology it is 2003-2018, and for the government's policy posture on innovation it is 2006-2018. All of the regressions include city and year-fixed effects (FE). Control variables include the logs of GDP, population and industry revenue the year before the local leader was appointed, and FEs for the term year of the provincial governor and party secretary. The initial conditions of the dependent variable include the log value and the growth rate in the 5 years preceding the appointment of the local leaders. Standard errors (SE) are clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Sci & Tech: science and technology; SD: standard deviation.

However, I also find that connected leaders are less engaged in promoting science and technology and innovation (row (1), columns (3)-(6)). When leaders are connected, the local government's budget spending on science and technology is lower by 10% (row(1), column (3)), which is equivalent to a reduction of 10.4 million RMB.[39] Moreover, local governments mention innovation less frequently in their GWRs by 0.074 SD (row (1), column (5)). The reduction in spending and policy posture could have different implications for firms and universities. As explained in section 2, direct financing from governments is crucial for innovation by universities due to their heavy financial dependence on government funding. For firms, both reductions can be consequential as firms' innovation activities are sensitive to both direct financing and indirect policy support. In the next section, I will explore the heterogeneous response in innovation by firms and universities in detail.

The contrast between higher infrastructure spending on the one hand, and lower spending on science and lower policy posture on innovation on the other hand implies a clear priority

[39] The average budget spending on science and technology by the city government is approximately 105 million RMB.

of fast over slow. Local governments' spending on infrastructure primarily goes to public infrastructures, such as water, gas, electricity, internet, telecommunication, and transportation. Aside from broad spillover effects in the future, however, infrastructure spending can also immediately stimulate investment, employment, and thus GDP growth.[40] However, the local governments' spending on science and technology primarily supports R&D activities through research grants and subsidies. And the yields from these investments arise much slower and are unlikely to immediately contribute to GDP growth.

I argue that the observed priority given to fast over slow policies is likely to be driven by leaders' turnover expectations. One natural concern is that connections can affect policy choices through other channels. For example, connected leaders generally receive more support from the provincial government to coordinate local development. However, I find that connections affect resource transfers neither explicitly nor implicitly. Another possibility is that connected leaders are more likely to seek rent through infrastructure development. However, I show that connections sharpen leaders' career concerns by enhancing their promotion prospects. In other words, connected leaders are less likely to be corrupted by trading future political careers for short-term economic benefit.[41]

5.3 Growth and innovation

I next examine the effects of political connection on growth and innovation based on equation (2). To relate this to the results on fast and slow policies on growth, I examine the impulse response of growth and innovation to connection. Table 5 presents the results on current growth (innovation) and future growth (innovation). Here the focus is on how connections affect the dynamics of growth and innovation.

Connected leaders deliver higher GDP growth in the short run but lower growth in the long run (row (1) in Panel A). The current effect of connection on GDP growth is 1 pp (row (1), column (1), Panel A), which is equivalent to an 8% increase relative to the average growth rate, or approximately 320 RMB in GDP per capita. The effect on GDP growth peaks in the following year (year t+1) (row(1), column (2), Panel A)), i.e. when leaders are

[40] 20% of physical investment go toward infrastructure development. 45% of infrastructure development is public infrastructure development which is mainly financed by the government.

[41] The corruption rate among city leaders is relatively low, and the connection rate is quite well-balanced between connected ones and unconnected ones. From 1994 to 2020, only approximately 4% of the city leaders were ever demoted or suspended due to corruption (calculated using the author's data). Moreover, among those who were later demoted/suspended, I do not find the leaders to have more connections than average. Therefore, although I cannot entirely rule out the possibility of connected leaders pursuing rent-seeking activities, I infer that rent-seeking driven by economic incentives is very limited among city leaders.

Table 5: Estimates of the effects of connections on growth and innovation

	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Grow Rate in GDP											
Connected	0.0102***	0.0189***	0.0159***	0.0101***	0.0042	-0.0002	-0.0096**	-0.0107***	-0.0064	0.0017	0.0041
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
$Connected^{start}$	-0.0025	-0.0125***	-0.0129***	-0.0056	-0.0059	0.0011	0.0153***	0.0121***	0.0073	0.0015	-0.0076
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Observations	4,419	4,156	3,885	3,618	3,340	3,062	2,814	2,539	2,267	2,004	1,731
R-squared	0.488	0.506	0.512	0.521	0.529	0.542	0.573	0.582	0.580	0.569	0.562
Mean	0.127	0.126	0.124	0.122	0.119	0.116	0.110	0.105	0.098	0.092	0.086
Panel B: log(# of invention patents applications)											
Connected	-0.1190***	-0.0972***	-0.0685**	-0.0481	-0.0614*	-0.0743**	-0.0798**	-0.0331	-0.0429	-0.0164	-0.0121
	(0.028)	(0.030)	(0.032)	(0.033)	(0.034)	(0.034)	(0.033)	(0.030)	(0.030)	(0.031)	(0.030)
$Connected^{start}$	0.0224	-0.0119	-0.0524	-0.0736*	-0.0637	-0.0484	-0.0247	-0.0103	-0.0148	-0.0135	0.0024
	(0.039)	(0.040)	(0.042)	(0.044)	(0.044)	(0.045)	(0.040)	(0.036)	(0.034)	(0.033)	(0.034)
Observations	4,874	4,590	4,326	4,056	3,782	3,502	3,252	2,973	2,696	2,420	2,141
R-squared	0.941	0.937	0.936	0.935	0.936	0.940	0.942	0.947	0.950	0.954	0.959
Mean	5.079	5.192	5.320	5.445	5.571	5.707	5.850	5.978	6.112	6.246	6.377
City and Year FE	X	X	X	X	X	X	X	X	X	X	X
Controls	X	X	X	X	X	X	X	X	X	X	X
Trend.init.Dep	X	X	X	X	X	X	X	X	X	X	X
SE Cluster	City	City	City	City	City	City	City	City	City	City	City

The unit of observation is city by year and the sample period is 2000-2018. The dependent variable in panel A is annual GDP growth and in panel B it is the log value of the total number of invention patent applications. Each column indicates the period of the corresponding outcome variable. All regressions include city and year fixed effects (FE). Control variables include the logs of GDP, population and industry revenue for the year before the local leader was appointed, and FEs for the term year of the provincial governor and party secretary. Standard errors (SE) are clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Trend.init.Dep: initial value of the outcome variable before the leader was appointed and the trend in the outcome variable 5 years before the leader was appointed.

connected today the GDP growth rate is higher by 1.9 pp in the next year. However, from the third year onwards, the effects of connections first decay, with the positive effects only lasting until the fourth year (year t+3) (row (1), column (4)), become negligible from then till the sixth year (row (1) columns (5)-(6), and flipped as negative for the seventh year and the eighth year (row (1), columns (7)-(8), with the GDP growth rate lower by 1 pp. The short-term effect on economic growth is similar to those reported in previous studies, which shows that connections sharpen leaders' incentives in economic work (Jia et al., 2015; Jiang, 2018). However, given the effect of connections on future growth, it is clear that connections only sharpen leaders' incentives for short-term economic growth.

In contrast, connected leaders reduce innovation in the future (row (1) in Panel B). I observe an immediate drop of 11.9% in the invention patent applications (row (1), column (1), Panel B), which is equivalent to 19 fewer patents in the current year. This effect persists over time and lasts at least until the seventh year (year t+6), with a magnitude of around 16-17 fewer patents in each corresponding year. These very dynamic effects highlight the slow-to-yield nature of innovation activities and their vulnerability to short-termism policies.

This contrast between growth and innovation is consistent with previous results on tenure expectation and policy priority of fast over slow. Fast policies such as infrastructure development can immediately boost investment and employment, have broad spillover effects on the local economy and facilitate innovation over time. However, local leaders pursue fast policies at the cost of innovation by spending less and working less on innovation. When the effect of underinvestment in innovation materializes, it supersedes the spillover effects of infrastructure and reduces growth.

Furthermore, my findings also suggest that both policies and market response might be at play. The effects on economic growth and innovation are considerably large, compared with the moderate effects on tenure expectation and policy choices. One possible explanation is that politicians' short-termism also changes the market environment and market expectations, which might amplify the distortion caused by short-termist policies. The dynamics in future patenting activity, particularly the immediate drop and the later persistent decline, is consistent with this explanation, as existing studies have shown that patent applications respond immediately to direct policy shocks but much more slowly to market shocks (Aghion et al., 2018, 2019).[42] I infer that the immediate drop is likely to be primarily driven

[42]From the perspective of a direct policy shock, a close comparison is Aghion et al. (2019), a study that exploits the state composition of the US Senate Appropriation Committee as the instrumental variable for innovation. They observe a sharp rise in patents applied by universities immediately after a one-member increase in state representation on the Appropriation Committee and a rise in patent applications by all sectors 3 years later. From the perspective of change in the market environment, a close comparison is

by a reduction in funding support that discourages the filing of patent applications and the persistent decline later is likely to be driven by both factors. While a full analysis to disentangle these two mechanisms is beyond the scope of this study, I will discuss the relative magnitudes of the direct effects and the indirect effects by comparing innovation by firms and universities in the next section.

6 Interpretation and Extension

The previous section shows that leaders' connections affect their turnover, policy choices, and innovation outcomes. I also show that cities with connected leaders are not significantly different from those with unconnected leaders and that connections do not affect the resource transfer from the provincial government. However, as connections affect both tenure and promotion, it remains unclear how the interplay between these factors affects the policy and innovation outcome.

I leverage the heterogeneous effects of connections by leaders' heterogeneity on innovation by different sectors to isolate the horizon effect from the promotion effect. Specifically, I study the heterogeneous effects of connections by leaders' age category and compare sub-samples of leaders whose tenure alone is affected by sub-samples of leaders whose tenure and promotion are both affected. I also examine the heterogeneous effects of connections on innovations by firms and universities, as the former are affected by both policies and market whereas the latter are primarily affected only by innovation policies.

6.1 Horizon Effect v.s. Promotion Incentive

I first study the heterogeneous effect by age group and report the results in Table 6. A leader is considered young if her age upon appointment is below the median age of the appointment cohort. The coefficient of *Connected* represents the effect of connections for young leaders, and the sum of *Connected* and $Connected * Old$ represents the effect of connections for old leaders.

I find that although the effect of connection on shortening tenure does not differ across age categories (rows (1)-(2), column (1)), the effect of connections on enhancing the promotion likelihood is only observed for old leaders (rows (1)-(2), column (2)). In other words, connections help young leaders move faster but assist old leaders in moving *up* faster. This

Aghion et al. (2018), a study that quantifies the impact of a demand shock on firm innovation in France and finds that a patent response arises 3-5 years later.

differential effect across age groups offers a unique opportunity to examine the relative magnitude of horizon and promotion incentives on outcome variables. First of all, conditional on connection, this approach can address the concerns of exclusion restriction due to connection affecting other unobservable factors that affect a leader's behaviors. Second, by leveraging the heterogeneous effect of the connections between old and young leaders, I can isolate the effect of promotion incentives.

Table 6: Estimates of the effects of connections by age group

	Turnover Outcome (dummy)		log(Gov Spending) (log)		Policy Posture (SD)
	Exited	Promoted	Infrastructure	Sci&Tech	Innovation
Variables	(1)	(2)	(3)	(4)	(5)
Connected	0.0467***	0.0124	0.2107***	-0.1480***	-0.0918
	(0.015)	(0.011)	(0.077)	(0.052)	(0.063)
*Connected*Old*	-0.0173	0.0392***	-0.1988**	0.0891	0.0236
	(0.020)	(0.015)	(0.100)	(0.061)	(0.071)
$Connected^{start}$	-0.0385**	-0.0182	-0.4254***	0.0811	0.1078
	(0.016)	(0.011)	(0.119)	(0.090)	(0.093)
$Connected^{start}*Old$	-0.0201	-0.0447***	0.3829***	-0.0696	-0.1543
	(0.021)	(0.016)	(0.142)	(0.104)	(0.105)
Old	0.0607***	-0.0338***	-0.1706	-0.0197	0.0911
	(0.013)	(0.009)	(0.121)	(0.083)	(0.092)
Observations	11,727	11,727	1,126	4,262	3,034
R-squared	0.285	0.123	0.896	0.935	0.712
Mean	0.245	0.098	6.402	4.581	-0.010
City and year FE	X	X	X	X	X
Controls	X	X	X	X	X
Init.Cond.Depvar			X	X	X
SE Cluster	City	City	City	City	City

For columns (1)-(2), the unit of observation is leader by year and the sample period is for leaders who are appointed during the period 1997-2017, including the mayor and the party secretary. For columns (2)-(5), the unit of observation is city by year and the sample period varies based on data availability. The dependent variables are an indicator of whether the leader leaves office in the given year (column (1)), whether she/he is promoted in the given year (column (2)), the log value of budget spending on infrastructure development (column (3)), the log value of budget spending on science and technology (column (4)), and government's policy posture on innovation (column (5)). The sample period for budget spending on infrastructure development is 2006-2016, for budget spending on science and technology it is 2003-2018, and for the government's policy posture on innovation it is 2006-2018. In columns (1) and column (2), term year fixed fixed effects are controled for. In columns (3)-(5), initial conditions of the outcome variable before the leader arrives are controled for (Init.Cond.Depvar). All regressions include city fixed effects and year fixed effects. Control variables include the logs of GDP, population and industry revenue for the year before the local leader was appointed, and fixed effects for the term year of the provincial governor and party secretary. The initial conditions of the dependent variable include the log value and the growth rate in the 5 years preceding the appointment of the local leader. Standard errors (SE) are clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Gov.: government; Sci & Tech: science and technology

I next examine whether the effect on economic policies also differs across age groups. I find that the policy pursuit of fast-over-slow is primarily driven by young leaders. When cities have young connected leaders, the local government's spending on infrastructure is

21% higher (row (1) column (3)), which amounts to a 126 million RMB increase, but the local government's spending on science and technology is 14.8% lower, which is equivalent to cuts of 12.6 million RMB. Similarly, cities with connected young leaders have lower policy postures on innovation by 0.09 SD (insignificant) (row (1), column (5)). The pursuit of fast policy over slow policy is much less prominent among old connected leaders (row (1) + row (2)). One possible explanation is that promotion incentives mitigate the problem of short horizons as old-connected leaders also have a higher likelihood of promotion.

6.2 Firms v.s. Universities

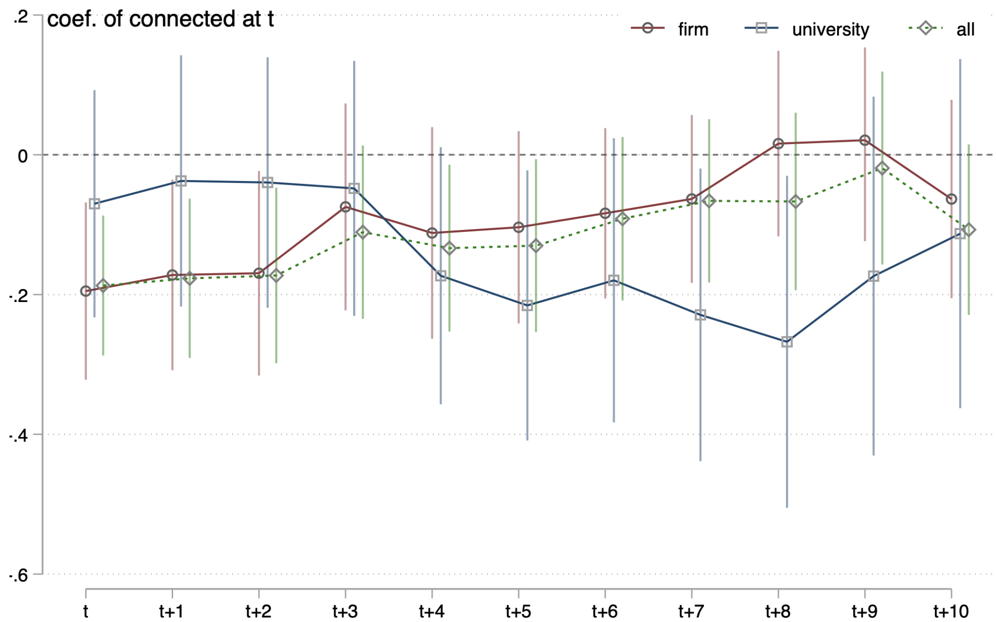


Figure 3: Dynamic effects of connection on innovation: firm v.s. university

The y-axis represents the point estimate and confidence interval of the estimated coefficient for the variable of *Connected* for each period, using the following specification:

$$
\begin{aligned}
y_{c,t+k} = \theta^{k} * Connected_{c,t} + \theta^{k,start} * Connected_{c,t}^{start} \\
\theta_{old}^{k} * Connected_{c,t} * Old_{c,t} + \theta_{old}^{k,start} * Connected_{c,t}^{start} * Old_{c,t} \\
+ X_{c,t}\Theta + \delta_c + \tau_t + \epsilon_{c,t}
\end{aligned}
$$

in which k ranges from 1 to 10.

To reconcile the findings on moderate tenure reduction with a considerable impact on the real economy, I examine the effects of political connections on innovation by firms and universities. They are the two largest institutional applicants in China, with the former making 60% and the latter making 15% of the patent applications, but differ in funding source and incentives for R&D activities.[43] While Chinese universities are primarily funded by the government and their innovation activities are likely to be affected directly by changes in direct funding support from the government, Chinese firms rely little on the government's direct financing but are highly responsive to indirect government financing, such as tax incentives, credit assistance, talent recruitment, patent services, and collaborations with research institutes.[44] A comparison between these two can help shed light upon how changes in policies affect growth and innovation. Using the same specification as in Table 5 but interacting the connection dummy with the age dummy, I compare how the patent responses of firms and universities differ over time and plot the estimated coefficients of *Connected* in Figure 3.

I observe differential dynamics in firms' and universities' patenting responses. For firms, I observe an immediate and persistent reduction in patent applications, with the effects decaying over time. When a city is exposed to a leader with a shorter horizon due to being connected, its local firms will reduce patenting activities by almost 20% in the following three years. From the third year and onwards, local firms' patenting activities are around 12% lower and this gap lasts till the seventh year. This finding is consistent with the results in section 5.3 and suggests that the aggregate change in patenting activities is primarily driven by firms. For universities, however, I find that the decline in patenting materializes later but last longer. In the first three years when local firms experienced a sharp decline in patenting activities, local universities' patenting activities barely change. From the fourth year and onwards, local universities' patenting activities shrink by up to 25% and the reduction last tills the ninth year.

6.3 Quantifying the Horizon Effect on Innovation

What is the effect on innovation if a leader's expected tenure is shorter by 1 year? I examine the effect of connection on patent applications by firms and universities from now until 5 years hence. Table 7 reports the results, and I draw information from columns (2) and column (4) for interpretation. In all of the regressions, *Connected* represents the semi-elasticity of innovation with regard to having young connected leaders.

[43] The remaining categories are by individuals (19%) and governments' affiliated research institutes (6%). The categories of individuals are relatively noisy as they could be either individuals or small firms (Sun et al., 2021), and it is unclear how market forces affect them. I also exclude the category of governments' affiliated research institutes as it is relatively small and, in practice, its orientation is between for-profit and not-for-profit.

[44] According to the survey by NBS in 2010, universities receive 56% of their R&D funding from the government. At the aggregate level, only 4% of the R&D expenditure by firms comes from the government in China

Table 7: Estimates of the effect of connection on the future number of patents

	Firm			University		
	All	L-impact	H-impact	All	L-impact	H-impact
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Connected	-0.1235*	-0.1278*	-0.1695**	-0.2474**	-0.2172**	-0.2354**
	(0.066)	(0.067)	(0.081)	(0.097)	(0.096)	(0.101)
*Connected*Old*	0.0436	0.0484	0.0710	0.1091	0.0811	0.0932
	(0.070)	(0.070)	(0.082)	(0.111)	(0.110)	(0.115)
$Connected^{start}$	-0.0765	-0.0794	-0.0296	0.1114	0.0944	0.1198
	(0.093)	(0.094)	(0.096)	(0.137)	(0.136)	(0.140)
$Connected^{start}*Old$	0.0302	0.0291	0.0152	-0.1415	-0.1243	-0.1360
	(0.101)	(0.101)	(0.101)	(0.152)	(0.151)	(0.153)
Old	-0.0565	-0.0577	-0.0669	0.0930	0.0986	0.0579
	(0.082)	(0.082)	(0.082)	(0.128)	(0.126)	(0.137)
Observations	3,478	3,478	3,478	3,478	3,478	3,478
R-squared	0.961	0.959	0.960	0.944	0.942	0.919
Mean	4.398	4.281	2.174	1.931	1.823	0.452
City and year FE	X	X	X	X	X	X
Controls	X	X	X	X	X	X
Init.Cond.Depvar	X	X	X	X	X	X
SE Cluster	City	City	City	City	City	City

The unit of observation is city by year and the sample period is from 2000-2018. The dependent variable is the log value of invention patent application by firms from years t to t+5 (columns (1)-(3)), and invention patent applications by universities from years t to t+5 (columns (4)-(6)). All regressions include city and year fixed effects (FE). Control variables include the logs of GDP, population and industry revenue for the year before the local leader was appointed, and FEs for the term year of the provincial governor and the party secretary. The initial conditions of the dependent variable (Init.Cond.Depvar) include the log value and the growth rate in the 5 years preceding the appointment of the local leader. Standard errors (SE) are clustered by city, in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The horizon semi-elasticity of innovation is sizable. When leaders expect a tenure shorter by 6 months (Table 6, row (1), column (1)), the annual innovation by local firms will be lower by 12.35% from now until five years later (row (1), column (2)). Following a simple back-of-the-envelope calculation, a one-year reduction in tenure implies a 27% drop in firms' innovation. For an average city, this is equivalent to 9 fewer patents per year. This effect is likely driven by both the direct and indirect effects of the fast-over-slow policy choice. By analogy, the effect on innovation by local universities is lower by 24.7% from now until 5 years later (row (1) column (3)). It implies that a one-year reduction in tenure can lead to a 55.4% reduction in universities' innovation. For an average city, this is equivalent to 3.4 fewer patents by the university annually. Furthermore, the horizon effect is similar between low-impact patents (row (1), column (2) and column (5)) and high-impact patents (row (1), column (3), and column (6)), thus excluding the possibility of quality-over-quantity in innovation policies.

7 Conclusion

Innovation is the key to addressing many of the unprecedented challenges in our societies — such as the global pandemic and climate change — and the only path to sustainable long-run productivity growth. Despite promoting innovation has become a central economic policy issue for many governments around the world, the questions of whether politicians should and could be held accountable have remained unanswered.

My paper contributes to answering these questions by highlighting the tension between the horizons politicians have and the horizons innovation needs. I examine how local politicians' tenure expectation affects their policy choices between promoting short-run growth and innovation in China. I construct a novel dataset that links leaders' careers to policies and innovation outcomes for all Chinese prefecture-level cities during the period 2000-2018. For identification, I draw on evidence that political connections put leaders on a fast track and exploit political connections formed through previous work ties to generate exogenous variation in leaders' tenure expectations. I find that when leaders can rationally expect shorter tenure, they spend more on infrastructure but less on innovation and put less emphasis on innovation. Such prioritization leads to higher short-run growth at the cost of innovation and long-run growth.

My data and empirical setup are particularly relevant for understanding China's ongoing transition from "made-in-china" towards "innovated-in-China". By zooming in on the short-run versus long-run policy-making and the consequence at the city level, I reach findings that are in contrast with the widely held perception that China's political stability enables long-term economic policies. My findings suggest that China's traditional bureaucrats' incentive, featured with frequent political turnover combined with economic performance competition, might erect an institutional barrier to innovation-led growth.

The insight in this paper, namely political short-termism leads to under-investment in innovation, may be generalized to other settings where politicians have considerable discretion over economic policies, such as those in Besley and Case (1995) and Dal Bó and Rossi (2011). The magnitude and the exact policy instruments are likely to vary from China due to differences in political and market institutions. But a future comparison between different settings will help expand our knowledge about institutional intermediaries that might mitigate or amplify political short-termism.

This paper also opens venues for future research on political accountability for innovation and appropriate growth policy. First, it raises the question of how to shield innovation from politics. Because politicians come and go for good reasons in most societies, as that's what keeps democracy in check, it might not be viable to solve the short-horizon issue by making politicians stay longer(Dewatripont et al., 1999; Smart and Sturm, 2013; Aghion and Jackson, 2016). However, there could be other forms of institutional solutions, such as allocating the decision-making to public officials that are not subject to political competition(Maskin and Tirole, 2004; Alesina and Tabellini, 2007, 2008). Future work could therefore focus on whether such arrangements can mitigate the short-termism in promoting innovation. Second, my study also points to the differential role of infrastructure investment and innovation investment for economic growth at different stages. I show that sacrificing innovation for infrastructure while approaching the technology frontier undermines future growth. Future work on how and when different economies switched their investment priorities to innovation might shed new light on the ongoing discussion on appropriate growth policy.

References

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hémous, D. (2019). Innovation and top income inequality. *The Review of Economic Studies*, 86(1):1–45.

Aghion, P., Bergeaud, A., Lequien, M., and Melitz, M. J. (2018). The heterogeneous impact of market size on innovation: evidence from french firm-level exports. Technical report, National Bureau of Economic Research.

Aghion, P. and Howitt, P. (2006). Appropriate growth policy: A unifying framework. *Journal of the european Economic Association*, 4(2-3):269–314.

Aghion, P. and Jackson, M. O. (2016). Inducing leaders to take risky decisions: dismissal, tenure, and term limits. *American Economic Journal: Microeconomics*, 8(3):1–38.

Aghion, P., Van Reenen, J., and Zingales, L. (2013). Innovation and institutional ownership. *American economic review*, 103(1):277–304.

Alesina, A. (1988). Macroeconomics and politics. *NBER macroeconomics annual*, 3:13–52.

Alesina, A. and Perotti, R. (1996). Fiscal discipline and the budget process. *The American Economic Review*, 86(2):401–407.

Alesina, A. and Tabellini, G. (2007). Bureaucrats or politicians? part i: a single policy task. *American Economic Review*, 97(1):169–179.

Alesina, A. and Tabellini, G. (2008). Bureaucrats or politicians? part ii: Multiple policy tasks. *Journal of Public Economics*, 92(3-4):426–447.

Besley, T. and Case, A. (1995). Does electoral accountability affect economic policy choices? evidence from gubernatorial term limits. *The Quarterly Journal of Economics*, 110(3):769–798.

Blanchard, O. and Shleifer, A. (2001). Federalism with and without political centralization: China versus russia. *IMF staff papers*, 48(1):171–179.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, 33(3):163–84.

Campante, F. R., Chor, D., and Li, B. (2019). The political economy consequences of china's export slowdown. Technical report, National Bureau of Economic Research.

Dal Bó, E. and Rossi, M. A. (2011). Term length and the effort of politicians. *The Review of Economic Studies*, 78(4):1237–1263.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *The Review of Economic Studies*, 66(1):199–217.

Ding, X. and Li, J. (2015). *Incentives for innovation in China: building an innovative economy*. Routledge.

Fang, J., He, H., and Li, N. (2020). China's rising iq (innovation quotient) and growth: Firm-level evidence. *Journal of Development Economics*, 147:102561.

Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274–1311.

Fisman, R. (2001). Estimating the value of political connections. *American economic review*, 91(4):1095–1102.

Fisman, R., Shi, J., Wang, Y., and Wu, W. (2020). Social ties and the selection of china's political elite. *American Economic Review*, 110(6):1752–81.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–74.

González-Uribe, J. and Groen-Xu, M. (2017). Ceo contract horizon and innovation. *NBER Working Paper*.

Green, R., Liyanage, S., Pitsis, T., Scott-Kemis, D., and Agarwal, R. (2010). The oecd innovation strategy: Getting a head start on tomorrow.

Hu, A. G., Zhang, P., and Zhao, L. (2017). China as number one? evidence from china's most recent patenting surge. *Journal of Development Economics*, 124:107–119.

Huang, Y., Pagano, M., and Panizza, U. (2020). Local crowding-out in china. *The Journal of Finance*, 75(6):2855–2898.

Jia, R. (2017). Pollution for promotion. *21st Century China Center Research Paper*, (2017-05).

Jia, R., Kudamatsu, M., and Seim, D. (2015). Political selection in china: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4):631–668.

Jia, R. and Nie, H. (2017). Decentralization, collusion, and coal mine deaths. *Review of Economics and Statistics*, 99(1):105–118.

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

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2018). Measuring technological innovation over the long run. Technical report, National Bureau of Economic Research.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132(2):665–712.

König, M., Song, Z. M., Storesletten, K., and Zilibotti, F. (2020). From imitation to innovation: Where is all that chinese r&d going? Technical report, National Bureau of Economic Research.

Kou, C.-w. and Tsai, W.-H. (2014). "sprinting with small steps" towards promotion: solutions for the age dilemma in the ccp cadre appointment system. *The China Journal*, (71):153–171.

Lerner, J. and Wulf, J. (2007). Innovation and incentives: Evidence from corporate r&d. *the Review of Economics and Statistics*, 89(4):634–644.

Li, H. and Zhou, L.-A. (2005). Political turnover and economic performance: the incentive role of personnel control in china. *Journal of public economics*, 89(9-10):1743–1762.

Li, X. (2012). Behind the recent surge of chinese patenting: An institutional view. *Research policy*, 41(1):236–249.

Li, X. (2019). The costs of workplace favoritism: Evidence from promotions in chinese high schools. *Manuscript. Hong Kong University of Science and Technology.*

Lindbeck, A. (1976). Stabilization policy in open economies with endogenous politicians. *The American Economic Review*, 66(2):1–19.

Martinez, C. (2010). Insight into different types of patent families.

Maskin, E., Qian, Y., and Xu, C. (2000). Incentives, information, and organizational form. *The review of economic studies*, 67(2):359–378.

Maskin, E. and Tirole, J. (2004). The politician and the judge: Accountability in government. *American Economic Review*, 94(4):1034–1054.

Naughton, B. (2017). Is china socialist? *Journal of Economic Perspectives*, 31(1):3–24.

Pang, B., Keng, S., and Zhong, L. (2018). Sprinting with small steps: China's cadre management and authoritarian resilience. *The China Journal*, 80(1):68–93.

Porter, M. E. (1992). Capital disadvantage: America's failing capital investment system. *Harvard business review*, 70(5):65–82.

Qian, Y., Roland, G., and Xu, C. (2006). Coordination and experimentation in m-form and u-form organizations. *Journal of Political Economy*, 114(2):366–402.

Rogoff, K. (1990). Equilibrium political budget cycles. *The American Economic Review*, pages 21–36.

Shambaugh, J., Nunn, R., and Portman, B. (2017). Eleven facts about innovation and patents. *Economic facts-The Hamilton Project*, 1.

Shi, X. and Xi, T. (2018). Race to safety: Political competition, neighborhood effects, and coal mine deaths in china. *Journal of Development Economics*, 131:79–95.

Shih, V., Adolph, C., Liu, M., et al. (2012). Getting ahead in the communist party: Explaining the advancement of central committee members in china. *American Political Science Review*, 106(01):166–187.

Shih, V. and Lee, J. (2020). Locking in fair weather friends: Assessing the fate of chinese communist elite when their patrons fall from power. *Party Politics*, 26(5):628–639.

Smart, M. and Sturm, D. M. (2013). Term limits and electoral accountability. *Journal of public economics*, 107:93–102.

Sun, Z., Lei, Z., Wright, B. D., Cohen, M., and Liu, T. (2021). Government targets, end-of-year patenting rush and innovative performance in china. *Nature biotechnology*, 39(9):1068–1075.

Titiunik, R. (2016). Drawing your senator from a jar: Term length and legislative behavior. *Political Science Research and Methods*, 4(2):293–316.

Wei, S.-J., Xie, Z., and Zhang, X. (2017). From" made in china" to" innovated in china": Necessity, prospect, and challenges. *Journal of Economic Perspectives*, 31(1):49–70.

Wong, S. H.-W. and Zeng, Y. (2018). Getting ahead by getting on the right track: Horizontal mobility in china's political selection process. *Journal of Contemporary China*, 27(109):61–84.

Xiong, W. (2018). The mandarin model of growth. Technical report, National Bureau of Economic Research.

Xu, C. (2011). The fundamental institutions of china's reforms and development. *Journal of economic literature*, 49(4):1076–1151.

Xu, G. (2018). The costs of patronage: Evidence from the british empire. *American Economic Review*, 108(11):3170–98.

Appendices

Appendix A Encoding CV Based on the Bureaucratic Hierarchy

The raw documents contain rich but unstructured information that cannot be analyzed directly. Each raw CV is a chronological account of events, in which each event, represented as a string, refers to a person working at one or several jobs for a certain period.[45]

I structure CV trajectories by encoding positions contained in CVs based on the existing Chinese bureaucratic hierarchy. I first parse each event string as a set of job vectors using the NLP-NER method. Next, I convert each person's CV into a year-job panel, with the job being defined in step one.

A.1 Definitions of Position

For each job string (contained in an event string), I apply the NER method to parse it as a set of positions, with each position defined as a tuple of locality, organization, and job title. This encoding approach creates 16,182 unique positions.

$$Position = (Locaility, Organization, JobTitle)$$

- The universe of *Locality* contains 3,360 elements, with each element referring to either the whole nation (if it is a national level position) (1 element), a province (31 elements), a prefecture (334 elements), or a county (2,933 elements).
- The universe of $Organization$ contains 129 elements, with each element referring to either the military, regional-level party committee, government, local people's congress, local people's political consultative conference committee, or a division under any of the above organizations such as the department of propaganda within the party committee, or the committee of development and reform within the government.
- The universe of $JobTitle$ contains 9 elements, with each element referring to either the head (vice-head), the division head (vice-division-head), the section head (vice-section head), or the section member.
- Each position also has a rank that is jointly determined by the locality, organization and job title[46]. There are 11 levels of rank in the bureaucratic hierarchy, with section

[45]for example, http://ldzl.people.com.cn/dfzlk/front/personPage15093.htm.

[46]for reference, https://en.wikipedia.org/wiki/Civil_Service_of_the_People%27s_Republic_of_China.

members who are grass-root level bureaucrats as the bottom layer and the national head as the top layer.

For example, the vice chairman of the Standing Committee of the People's Congress of Xiamen Prefecture is a position with the following three attributes: $Locality$ =Xiamen, $Organization$ =local people's congress, $JobTitle$ = vice-head. Based on the ranking rules, this position is ranked as a deputy bureau director.

A.2 Definition of Superior-subordinate Relationship Between Positions

I denote the hierarchy of position or position features (location, organization or job title) as a graph $G = (N, E)$, in which N represents the set of nodes and E represents the set of directed edges in which each edge points from a superior to a subordinate. For any $n \in N$, its superior nodes ($sup(n)$ and weak superior ($wsup(n)$) nodes are defined as

$$sup(n) = \{n' \in N | (n', n) \in E\} \tag{A.1}$$

$$wsup(n) = \{n' \in N | (n', n) \in E\} \cup \{n\} \tag{A.2}$$

By this definition, the superior nodes of a node only refers to direct superiors, whereas weak superiors do not only refer direct superiors but also the node itself.

There exist three different types of superior-subordinate relationships in the bureaucratic system in China. The first is within-location-within-organization, such as the head of the anti-corruption bureau and the division head within the anti-corruption bureau of the same region. The second is within-region-between-ogranization, such as the anti-corruption bureau being supervised by the procuratorate of the same region. The third is between-region-within-organization, such as the anti-corruption bureau of the province of Hunan overseeing the anti-corruption bureau of its provincial capital city.[47]

It is worth emphasizing that the superior-subordinate relationship between organizations is restricted to only the following two situations:

- departments within a local government/party supervised by the local government head(vice-head) and the local party committee. For example, the head of government (mayor) is the direct superior of all department heads in the local government.

[47]Since 1984, the management of cadres has followed two rules: appointment made by the party committee one level up, and are subject to the leadership of both the local authority (in local government or in the local party committee) and the same department at an upper level.

- department-affiliated organizations supervised by the department to which they are affiliated. For example, the anti-corruption bureau is supervised by the procuratorate in the same region.

Under this institutional set up, a position m is a direct superior of another position n if there exists a weak superior-subordinate relationship in location and organization, and a superior-subordinate relationship in job titles between m and n.

$$\begin{aligned} &1\{Position_m \in sup(Position_n)\} \\ &= (1\{Locaility_m \in wsup_{loc}(Locaility_n)\} \\ &* 1\{Organization_m \in wsup_{org}(Organization_n)\} \\ &* 1\{JobTitle_m \in sup_{title}(JobTitle_n)\}) \end{aligned} \tag{A.3}$$

This definition of the superior-subordinate relationship allows for the following scenarios in which n has a title rank that is one or two levels above m

- m and n are from the same location and the same organization. For example, the head of the anti-corruption bureau and the division head within the anti-corruption bureau.
- m and n are from the same location and different organizations, but the organization of n supervises the organization of m. For example, the anti-corruption bureau is supervised by the procuratorate in the same region.
- m and n are from different locations, but the same organizations, and n's organization supervises m's organization due to the hierarchical relationship between their locations. For example, the anti-corruption bureau of a province oversees that of its provincial capital city.

A.3 Algorithms to Create the Hierarchy of Positions

I denote the total number of positions as N. The superior-subordinate relationship between positions is the Hadamard product of three $N \times N$ matrices, defined as follows

$$H = S_{location} \odot S_{org} \odot S_{rank} \tag{A.4}$$

- H is an $N \times N$ matrix that represents the superior-subordinate relationship between the row and the column positions, with each element being either 0 or 1.

- $S_{location}$ is an $N \times N$ matrix that represents the superior-subordinate relationship between the row and column positions' localities, with each element being either 0 or 1.

- S_{org} is an $N \times N$ matrix that represents the superior-subordinate relationship between the row position's organization and the column position's organization, with each element as either 0 or 1.

- S_{rank} is an $N \times N$ matrix that represents the superior-subordinate relationship between the row and the column positions' job titles (rank), with each element being either 0 or 1.

- $\odot$ represents the Hadamard product (element-wise product) between the matrix on the left-hand side and the matrix on the right-hand side of this operator symbol.

The derivation of the matrix of S_{loc} requires the matrix that represents the geographic location of position and the matrix that represents the hierarchical relationship between locations. I denote the total number of localities as N_{loc}, and the matrix of S_{loc} as:

$$S_{loc} = L \cdot H_{loc} \cdot (L)' \tag{A.5}$$

- L is an $N \times N_{loc}$ matrix that represents the geographic location relationship between the row position and the column location, with each element being either 0 or 1.

- H_{loc} is an $N_{loc} \times N_{loc}$ matrix that represents the superior-subordinate relationship between the row and the column positions, with each element being either 0 or 1 and all of the elements on the diagonal being 1.

The derivation of the matrix of S_{org} requires the matrix that represents the organization of position and the matrix that represents the hierarchical relationship between organizations. I denote the total number of organizations as N_{org} and the matrix of S_{org} as:

$$S_{org} = G \cdot H_{org} \cdot (G)' \tag{A.6}$$

- G is an $N \times N_{org}$ matrix that represents the organizational relationship between the row position and the column organization, with each element being either 0 or 1.

- H_{org} is an $N_{org} \times N_{org}$ matrix that represents the superior-subordinate relationship between the row and the column positions, with each element being either 0 or 1 and all of the elements on the diagonal being 1.

The derivation of the matrix of S_{rank} requires the matrix that represents the rank of position and the matrix that represents the hierarchical relationship between ranks. I denote the total number of ranks as N_{rank} and the matrix of S_{rank} as:

$$S_{rank} = K \cdot H_{rank} \cdot (K)' \tag{A.7}$$

- K is an $N \times N_{rank}$ matrix that represents the rank relationship between the row position and the column rank, with each element being either 0 or 1.
- H_{rank} is an $N_{rank} \times N_{rank}$ matrix that represents the superior-subordinate relationship in rank between the row and the column positions, with each element being either 0 or 1 and all elements on the diagonal being 0.

Appendix B Measuring Connection Based on Career Trajectory

B.1 Algorithms to Create Superior-subordinate Links

To further restrict the superior-subordinate relationship to represent patron-client ties, I first impose the superior-before-subordinate relationship between superior and subordinate.

$$\xi(t) = D(t) \cdot B \cdot (D(t))' \tag{B.1}$$

- $\xi(t)$ is an $N_{psn} \times N_{psn}$ matrix and N_{psn} represents the total number of persons. An element equals to 1 if the row(person) starts her/his current position before the column (person), and 0 otherwise.
- $D(t)$ is an $N_{psn} \times N_{dn}$ matrix that represents the person-duration relationship between the row person and the column duration (measured in years) at time t, with each element being either 0 or 1 and $N_{dn} = 10$. For example, $D(t)[i, 2] = 1$ if at time t the individual i is in the second year of her/his current position.
- B is an $N_{dn} \times N_{dn}$ matrix that represents the before-after relationship between the row and the column duration (years). If the row duration is not smaller than the column duration, then the corresponding element equals 1, and 0 otherwise.

I then use a square matrix $S(t)$ to denote the superior-subordinate relationship between individuals at time t. The derivation of $S(t)$ relies on a matrix that represents the relationship between person and position, a matrix that represents the hierarchy of positions and the matrix $\xi(t)$ defined as above

$$S(t) = ((Position(t)) \cdot H \cdot (Position(t))') \odot \xi(t) \tag{B.2}$$

- $S(t)$ is an $N_{psn} \times N_{psn}$ matrix and N_{psn} represents the total number of individuals. An element equals 1 if the row (person) supervises the column (person), and 0 otherwise.
- $Position(t)$ is an $N_{psn} \times N$ matrix that represents the person-position relationship between the row person and the column position at time t, with each element being either 0 or 1.
- H is an $N \times N$ matrix that represents the superior-subordinate relationship between the row and the column positions, with each element being either 0 or 1, and it is derived from equation A.4.

This algorithm enables a speedy and comprehensive search through the career trajectory network for subordinate-superior links. For individual i at time t, her/his direct superiors are defined as individuals who were holding a position that supervises person i's position at time t, which is equivalent to the i-th column of the matrix $S(t)$

$$sup(i,t) = \{j \in I : Position_{j,t} \in sup(Position_{i,t})\} := S(t)[:,i] \tag{B.3}$$

For individual i at time t, her/his previous superiors are defined as individuals who were holding a position that supervises person i's position before or at time t, denoted as $\bigcup_{s<t} sup(i,s)$.

B.2 Evaluating Connection Between Politicians

For any superior $j \in \bigcup_{s<t} sup(i,s)$, the value of i connected to j via previous subordinate-superior links depends on two factors: the intensity of the link, and the current political rank of j. I denote the intensity as $w_{i,j,t}$, and the political power of j as $\kappa_{j,t}$, and model the value of i connected to j, denoted as $v_{i,j,t}$, as follows:

$$v_{i,j,t} = \omega_{i,j,t}\kappa_{j,t} \tag{B.4}$$

I infer the value of $\omega_{i,j,t}$ from the time composition of the links. I let $T_{i,j,t}$ denote the set of periods prior to t when j was the direct superior of i

$$T_{i,j,t} = \{s < t | Position_{j,s} \in sup(Position_{i,s})\} \tag{B.5}$$

I impose the following three monotonicity constraints on $\omega_{i,j,t}$:

- $\omega_{i,j,t}$ is non-decreasing in τ_{min} conditional on $||T_{i,j,t}||$: the more recently the links are formed, the higher the connection intensity
- $\omega_{i,j,t}$ is non-decreasing in τ_{max} conditional on τ_{min}: the longer the links last between i and j, the higher the connection intensity
- $\omega_{i,j,t}$ is non-increasing in τ_{min} conditional on τ_{max}: the earlier the links start between i and j, the higher the connection intensity

Among the eligible function forms, I consider the following choices:

- time-discounted (with a discount constant δ)
  - sum: $\omega_{i,j,t} = \left(\sum_{\tau \in T_{i,j,t}} \delta^{t-\tau} \right)$
  - maximum :$\omega_{i,j,t} = \max_{\tau \in T_{i,j,t}} ||T_{i,j,t}|| \times \delta^{t-\tau}$
- time-constant
  - count :$\omega_{i,j,t} = ||T_{i,j,t}||$
  - binary :$\omega_{i,j,t} = 1\{||T_{i,j,t}|| >= 1\}$

I take the maximum value of connection at the individual level in the presence of multiple previous superiors. I let $v_{i,t}$ denote the value of the political connection person i at time t,

$$v_{i,t} = \max_{j \in \bigcup_{s<t} sup(i,s)} v_{i,j,t} \tag{B.6}$$

B.3 Evaluating the Connection Between City Leaders and Provincial Leaders

For the main specification of this study, I choose the binary measure of link intensity and restrict the links to be formed before the latest appointment of the city leader and the provincial leaders. Specifically, for city leader i and provincial leader j, the periods of links are defined as follows:

$$T_{i,j,t_0} = \{s < t_0 | Position_{j,s} \in sup(Position_{i,s})\} \tag{B.7}$$

where t_0 indicates the time when i and j encounter each other as city leader and provincial leader for the first time. The connection dummy is defined as whether i ever worked with j as subordinate-superior before t_0

$$v_{i,j,t} = 1\{||T_{i,j,t_0}|| >= 1\} \tag{B.8}$$

For each individual i, she/he is connected to her/his provincial leaders if she/he is connected to at least one of them:

$$v_{i,t} = 1\{v_{i,gr(i),t} + v_{i,ps(i),t} >= 1\} \tag{B.9}$$

where $gr(i)$ indexes the provincial governor of i and $ps(i)$ indexes the provincial party secretary.

Appendix C Validating the Empirical Design

C.1 Whether Leaders Select into Connection

I test whether career background affects a leader's probability of beginning as connected or later becoming connected, using the following specification:

$$ConnEvent_{i,c} = X_i\Omega + \tau_{t_{i,c}} + \delta_c + \epsilon_{i,c} \tag{C.1}$$

where X_i is a set of dummies that characterize a leader's career background, including whether the starting age is above the median of the same appointment cohort, whether having studied STEM at the university or worked in a STEM related department in the government, whether on a fast-track, whether connected to any other leaders at the provincial level, and whether having worked as a city leader before. $\tau_{i,t_{i,c}}$ indexes the starting year fixed effect (FE) for the spell. δ_c indexes the city FE.

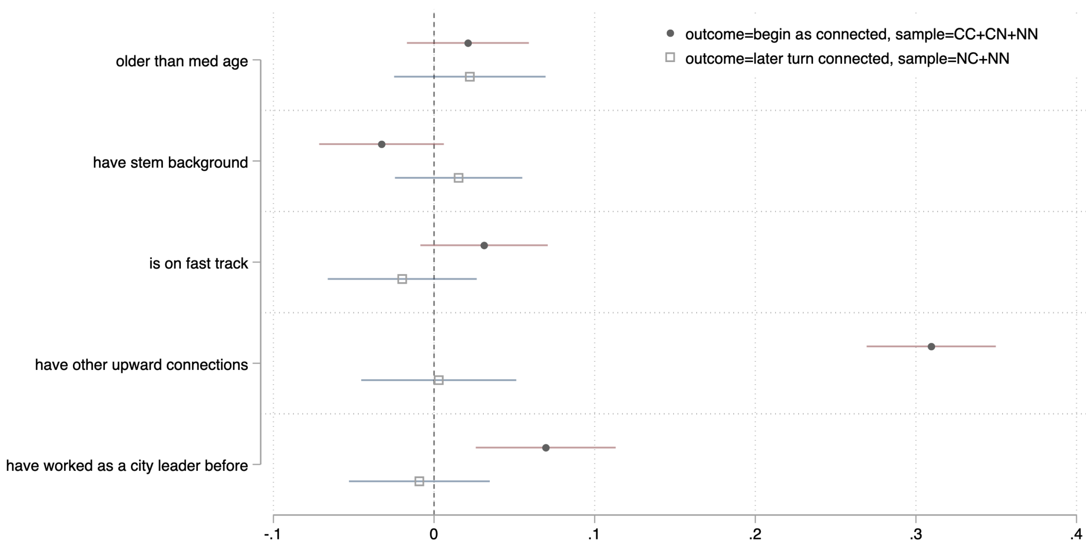


Figure C.1: Coefficients of leaders' background features on becoming connected

The unit of observation is the city-year spell for spells that started during the period 1997-2017. All regressions include FEs for the position (mayor or party secretary), and FEs for the term year of the provincial governor and the party secretary

C.2 Whether Cities Select into Having Connected Leaders

I test whether the initial social-economic conditions affect a city's probability of having a leader starting as connected or gradually becoming connected, using the following specification:

$$
\begin{aligned}
ConnEvent_{c,s} = {} & \eta_1' X^0_{c,s} + \eta_2' \Delta y^0_{c,s} \\
& + CityFE + TrendFE + unobservable
\end{aligned}
\tag{C.2}
$$

c index city and s index a tenure spell. $ConnEvent_{c,s} = 1$ if the leader either started as connected due to previous links or gradually became connected under the same provincial incumbents. $X^0_{c,s}$ is a set of variables to represent local conditions, including GDP, population, industry revenue (size of the manufacturing sector). $\Delta y^0_{c,s}$ is the 2-year-accumulated-change (or the 5-year-accumulated-change) in these variables preceding the appointment of the leaders to measure pre-trends. $TrendFE$ includes both year fixed effects in terms of the year in which the spell started, and the provincial leader's term year fixed effects. Standard errors are clustered by city.

Figure C.2 displays the results of the correlation between $X^0_{c,s}$ and connection events. Overall, cities' initial economic conditions are not correlated with the occurrence of various types of connection event, except for cities with smaller populations having a higher likelihood of being assigned leaders who begin as connected.

Figure D.3 displays the results of the correlation between $\Delta y^0_{c,s}$ and the connection events for each outcome variable, conditional on $X^0_{c,s}$.

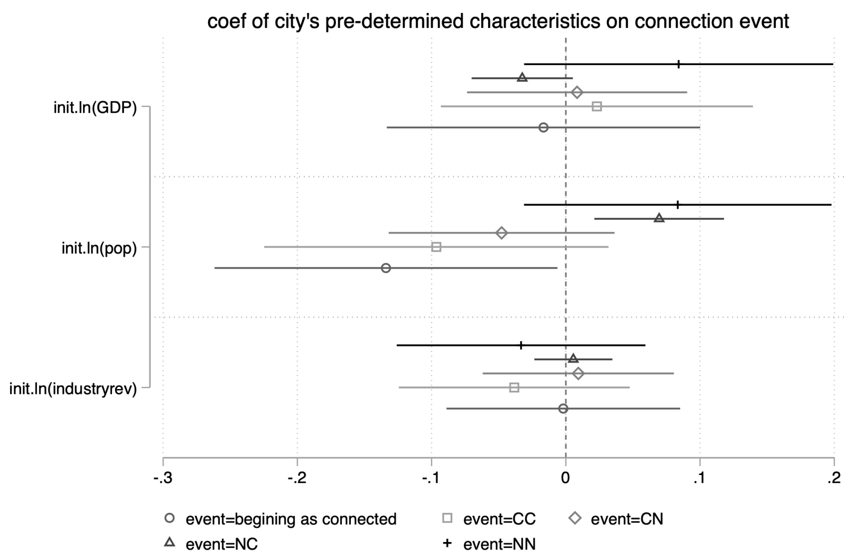
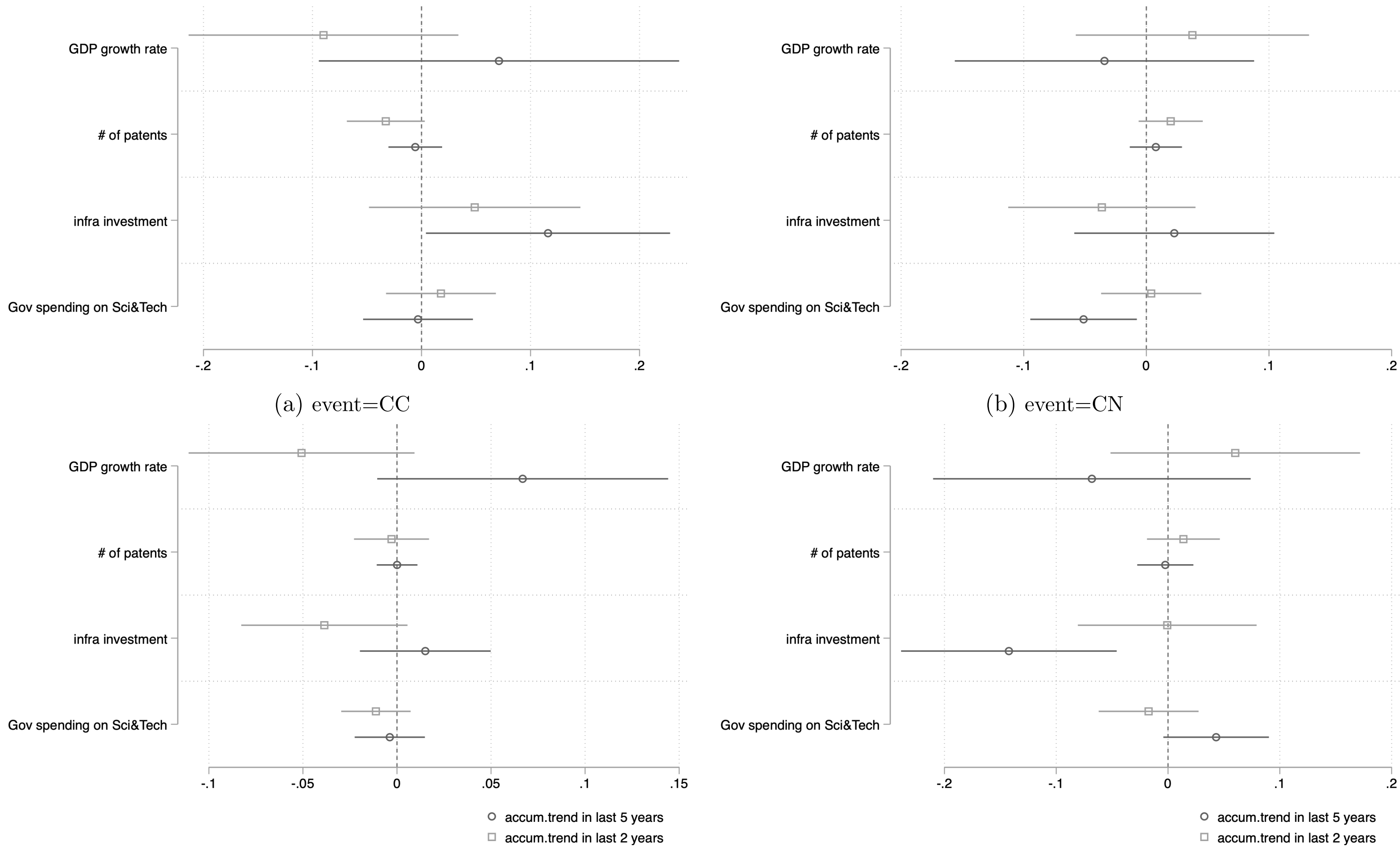


Figure C.2: Coefficients of the city's initial economic condition on connection events

The unit of observation is the city-year spell for spells that started during the period 1997-2017. All regressions include FEs for the position (mayor or party secretary), FEs for the term year of the provincial governor and the party secretary, and City FEs. Standard errors are clustered by city.



(a) event=CC

(b) event=CN

(c) event=NC

(d) event=NN

Figure C.3: Coefficients of the city's initial economic condition on connection events: outcomes

The unit of observation is the city-year spell for spells that started during the period 1997-2017. All regressions include FEs for the position (mayor or party secretary), FEs for the term year of the provincial governor and the party secretary, and City FEs. Standard errors are clustered by city.

C.3 Pre-trends in Patents and Growth

To test whether there exist pre-trends in innovation and growth for cities with connected leaders, I run the following regression:

$$
\begin{aligned}
y_{t-k} = {} & \theta_k * Connected_{c,t} + \theta_k^{start} * Connected_{c,t}^{start} \\
& + X_{c,t-k}\Theta_k + \delta_c + \tau_{t-k} + \epsilon_{c,t-k}
\end{aligned}
\tag{C.3}
$$

where k indicates the lag, ranging from 1 year to 5 years. For each lagged outcome, the control variables $X_{c,t-k}$ include GDP, population and manufacturing sector size 5 years before (and hence $t-k-5$), and an indicator of the provincial leader's term year at $t-k$. The fixed effects include city fixed effect δ_c and year fixed effect τ_{t-k}.

Table C.1 reports the results.

Table C.1: Estimates of the pre-trend of connection on growth and innovation

Variables	t-1 (1)	t-2 (2)	t-3 (3)	t-4 (4)	t-5 (5)
Panel A: Grow Rate in GDP					
Connected	-0.0023	-0.0075***	-0.0048	-0.0008	0.0015
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
$Connected^{start}$	0.0137***	0.0082*	-0.0009	-0.0045	-0.0017
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Observations	3,748	3,700	3,602	3,488	3,335
R-squared	0.547	0.539	0.521	0.468	0.429
Mean	0.133	0.138	0.143	0.152	0.159
Panel B: log(# of invention patents application)					
Connected	-0.0054	0.0093	0.0215	0.0092	0.0185
	(0.030)	(0.029)	(0.030)	(0.029)	(0.029)
$Connected^{start}$	-0.0383	-0.0056	-0.0111	0.0381	-0.0028
	(0.046)	(0.042)	(0.042)	(0.042)	(0.043)
Observations	3,753	3,704	3,608	3,501	3,349
R-squared	0.934	0.929	0.928	0.928	0.928
Mean	5.357	5.142	4.943	4.749	4.576
City and Year FE	X	X	X	X	X
Controls	X	X	X	X	X
SE Cluster	City	City	City	City	City

The unit of observation is city by year and the sample period is 2000-2018. The dependent variable in panel A is annual GDP growth and in panel B it is the log value of the total number of invention patent applications. Each column indicates the period of the corresponding outcome variable. All regressions include city fixed effects and year fixed effects. The control variables include the logs of GDP, population and industry revenue 5 years prior to t-k, and fixed effects for the term year of the provincial governor and the party secretary at t-k. Standard errors all clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Appendix D Robustness Checks

D.1 Regression on Tenure at the Spell Level

There exists a mechanical bias in linking tenure length to connection status directly. The longer a leader stays, the more likely that the provincial leaders change, which would decrease the probability for those beginning as connected to stay connected while increase the probability for those beginning as connected to gain connection. Figure D.1 illustrates the mechanical bias. Regressing tenure lengths on connection status directly will be biased by these two types of mechanical bias. One solution is to control for how long has leaders stay and run a year-by-year regression, as that in Section 4.2.

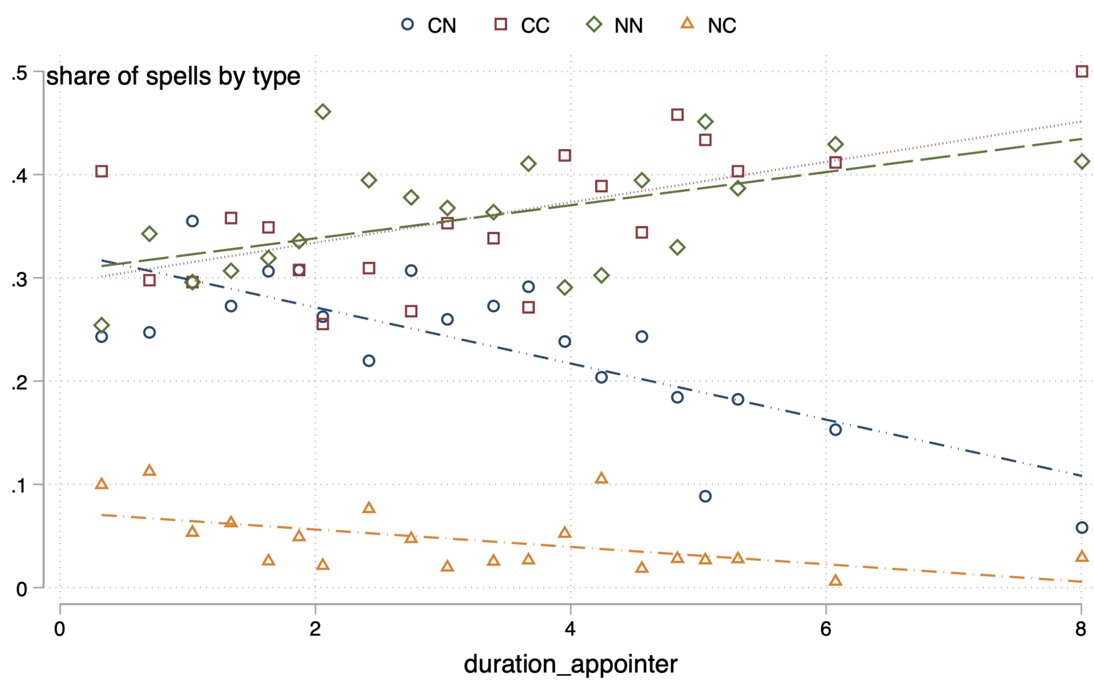


Figure D.1: Type composition of spell by superior's duration after the appointment

This plot presents the binscattered plot for the correlation between the probability of a city-leader spell experiencing a specific type of connection status transition and the duration of the appointers (i.e. provincial leaders who appointed the city leader in the first place) in years.

Another solution is to control for how long the provincial leaders stay after the appointment and run regression at the spell level. Specifically, I implement the following specification:

$$
\begin{aligned}
\ln(T_{i,c}) = {} & \delta * T_{i,c}^{sup} \\
& + (\delta * Connected_{i,c}^{end} + \eta * Connected_{i,c}^{start}) * T_{i,c}^{sup} \\
& + FEs + \nu_{i,c}
\end{aligned}
\tag{D.1}
$$

Conditional on the superiors' duration after appointing city leader i to city c (denoted as $T_{i,c}^{sup}$), the tenure for leader who is CC is then $(\delta + \eta) * \overline{T^{sup}}$., for leader who is NC is then $\delta * \overline{T^{sup}}$, for leader who is CN is then $\eta * \overline{T^{sup}}$, and the benchmark is leader who is NN. The intuition is that conditional on superiors' duration, the difference in tenure length is then driven by whether leaders are endogenously selected ($Connected_{i,c}^{end}$) and whether leaders are exposed to connection ($Connected_{i,c}^{start}$). For instance, for each additional year that the superiors stay after appointing a city leader, the difference between CC and CN is δ.

Table D.1 reports the results using the above specification. Overall, the results are similar with those using year-by-year estimation of the exiting rate.

Table D.1: Estimates of the effects of connection on tenure and promotion at spell level

	$ln(Tenure)$			Promoted		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
T^{sup}	0.0344***	0.0207***	0.0201***	0.0001	0.0023	0.0014
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
$T^{sup}*Connected^{end}$	-0.0814***	-0.0455***	-0.0453***	0.0088	0.0044	0.0056
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
$T^{sup}*Connected^{start}$	0.0586***	0.0310***	0.0315***	-0.0141**	-0.0099	-0.0107
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Old			-0.0330**			-0.2080***
			(0.016)			(0.019)
$FastTrack$			-0.0363*			0.0443**
			(0.019)			(0.021)
Observations	3,493	2,855	2,855	3,491	2,854	2,854
R-squared	0.257	0.490	0.491	0.179	0.198	0.239
Mean	1.092	1.074	1.074	0.391	0.393	0.393
City and year FE	X	X	X	X	X	X
Controls	.	X	X	.	X	X
SE Cluster	City	City	City	City	City	City

The unit of observation is leader-city spell. The sample is a pooled sample of leaders, namely city mayors and party secretaries, who was appointed during 2000-2017[a]. log value of term length (in years) (columns 1-3) or is promoted (columns 4-6). Controls include the logs of GDP, population and industry revenue the year before the local leader was appointed, and fixed effects for the term year of the provincial governor and party secretary. Standard errors (SE), clustered by city, in parentheses: *** p<0.01, ** p<0.05, * p<0.1..

[a]2018 is skipped due to data truncation issue, as around 20% of leaders are still in term

D.2 Event Study

At any time, a city is between the arrival of the current superiors and the arrival of the next superiors. Within each spell, for each superior leadership, the connection status is constant as it is pre-determined. The turnover of superiors hence creates a setting for event study. Figure D.2 illustrates this idea. For example, suppose a province and is governed by provincial leader team A for the period of 2003-2008 and leader team B for the period of 2009-2012. For a city leader a in year 2005, it is two years after A being in office and 3 years before B being in office. If a is connected to A but not connected to B, in 2005 a will have been connected to A for two years. If a is not connected to A but connected to B, in 2005 a is three years before becoming connected.

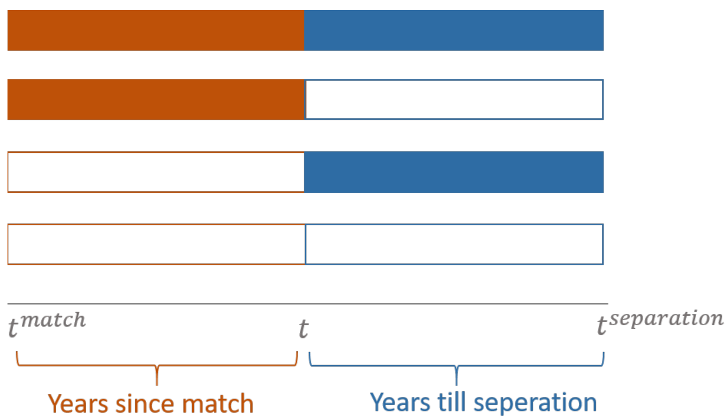


Figure D.2: Illustration of event study

In each bar, the filled color represents city leader being connected and empty color implies city leader being unconnected. The red color represents the connection status with the provincial incumbents. The blue color represents connection status with the next provincial leader.

I examine the before-after change of city's innovation outcome within each provincial leader spell

$$
\begin{aligned}
y_{c,t} = & \sum_{k=1}^{k=5} \theta_{-k} 1\{\text{k Yrs before } SupArrival^{next}\} * Connected_{c,t}^{next} \\
& + \sum_{k=1}^{k=5} \theta_k 1\{\text{k Yrs after } SupArrival\} * Connected_{c,t} + Controls + FEs + u_{c,t}
\end{aligned}
$$

Where $SupArrival$ indicates the year when the current provincial leadership arrives, and $SupArrival^{next}$ indicates the year when the current provincial leadership leaves. $Controls$ include the interaction terms between the after/till year dummies with $Connected^{started}$, and the control sets used in equation 2. FEs include city, term year and year fixed effects.

D.3 Other Type of Connections

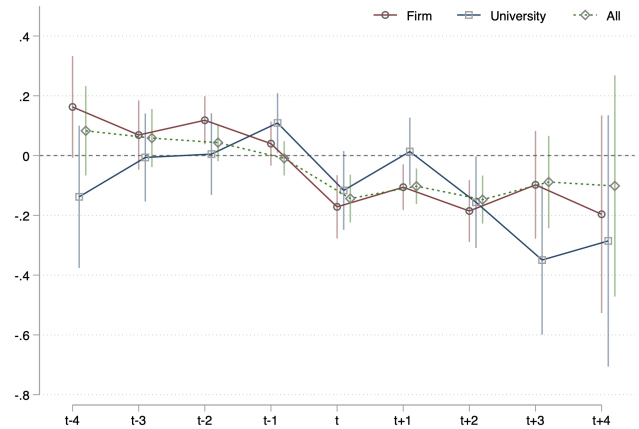
I examine whether connections formed through attending the same university or being from the same home town affect outcomes. Such types of connections are relatively sparse: at the spell level, 2.7% of city leaders are connected through university and 4% are connected through home town, and 6.5% of the spells are connected through either. Table D.2 presents the results following equation (1). Overall, I do not find these types of connections to be influential on leader's tenures.

Table D.2: Estimates of the effects of other type of connections on tenure and promotion

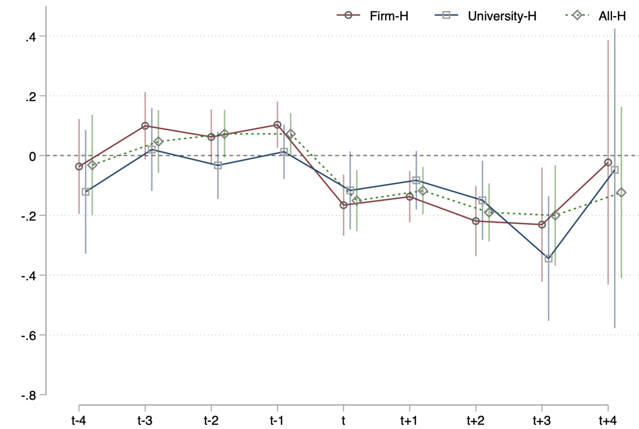
	Exited			Promoted		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$SameUniversity$	-0.0193			0.0209		
	(0.022)			(0.019)		
$SameHometown$		0.0025			0.0239	
		(0.019)			(0.017)	
$SameEither$			-0.0104			0.0159
			(0.015)			(0.013)
$STEM$	-0.0037	-0.0038	-0.0036	0.0087	0.0103*	0.0086
	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)
Old	0.0408***	0.0411***	0.0408***	-0.0317***	-0.0302***	-0.0318***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Observations	12,397	12,242	12,397	12,335	12,180	12,335
R-squared	0.274	0.252	0.273	0.107	0.106	0.107
Mean	0.094	0.092	0.094	0.094	0.092	0.094
City and year FE	X	X	X	X	X	X
Controls	X	X	X	X	X	X
SE Cluster	City	City	City	City	City	City

The unit of observation is leader by year. The sample is a pooled sample of leaders, namely city mayors and party secretaries, who were appointed during the period of 2000-2017[a]. The dependent variable is an indicator of whether the leader leaves office (columns (1)-(4)) or is promoted (columns (5)-(7)) in a given year. The variable $SameUniversity$ is a dummy and equals to 1 if the leader had attended the same university that the provincial leaders attended. The variable $SameHometown$ is a dummy and equals to 1 if the leader is from the same city where the povincial leaders are from. The variable $SameEither$ is a dummy and equals to 1 if either $SameUniversity$ or $SameHometown$ equals to 1. The variable *Old* is a dummy and equals to 1 if the leader's age upon appointment is older than the median age of all leaders appointed in the same year. The variable $STEM$ is a dummy that equals 1 if the leader has worked at a STEM related department in her/his previous jobs or has studied STEM at university. All of the regressions include a fully non-parametric baseline hazard for the number of years in office (job-tenure fixed effects (FE)). Column 1 shows results from a binomial regression with a complementary log-log model (marginal effects). The other columns show estimates using a linear probability model. Controls include the logs of GDP, population and industry revenue for year before the local leader was appointed, and FEs for the term year of the provincial governor and party secretary. Standard errors (SE) are clustered by city, in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

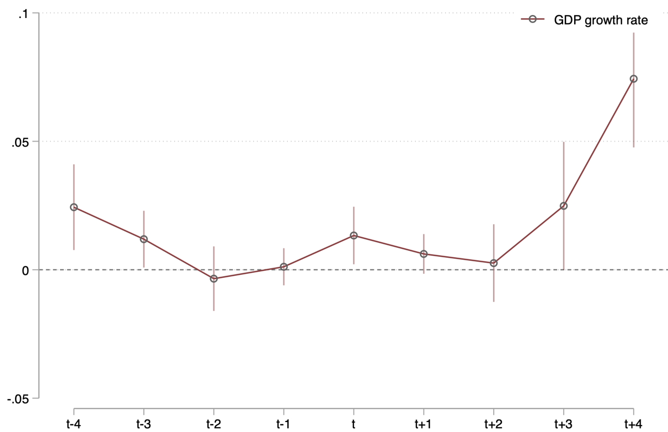
[a]2018 is skipped due to data truncation issues, as around 20% of the leaders appointed in 2018 still hold office.



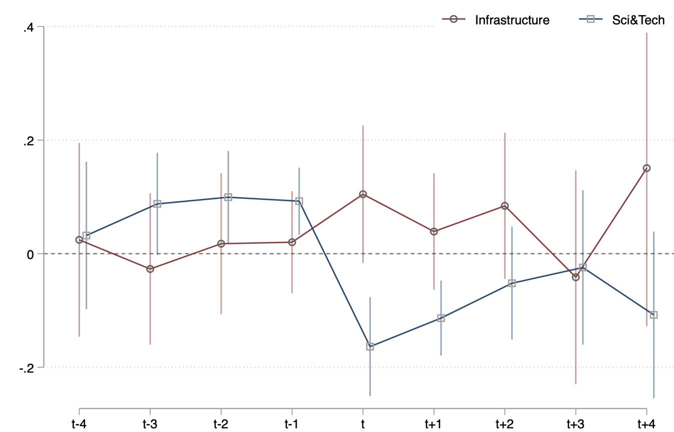
(a) log(# patents)



(b) log(# H- patents)



(c) GDP growth rate



(d) log(Gov.spending)

Figure D.3: Coefficients of *Connected* for event study

The unit of observation is city-year. For (a), (b) and (c) the sample period is 2000-2018 and all of the regressions control for the initial trend of the outcome variable preceding the appointment of city leaders. For (d), the sample period for government spending on science& technology is 2003-2018, and for government spending on infrastructure development it is 2006-2016.