Social Ties and Favoritism in Chinese Science

Raymond Fisman, Jing Shi, Yongxiang Wang, and Rong Xu*

This Draft: November 2015

Abstract

We study favoritism in the election of fellows to the Chinese Academies of Sciences and Engineering, appointments that confer considerable prestige and control over resources. We focus on hometown ties – a common source of favor-exchange in China – to department standing committees which organize the selection process. Hometown ties increase the probability of selection by 36 percent, which comes entirely from the selection stage involving an in-person meeting and vote. This effect disappears in 2007, likely owing to more stringent voting rules and greater transparency. We finally show that favoritism leads to lower-quality applicants gaining election to the CAS/CAE.

Keywords: knowledge production; social networks; resource misallocation

*Raymond Fisman: Economics Department, Boston University, 270 Bay State Road, Boston, MA; email: rfisman@bu.edu; Jing Shi: School of Economics, Finance and Marketing, RMIT College of Business, RMIT University, 445 Swanston Street Melbourne VIC 3000 Australia; email: jing.shi@rmit.edu.au; Yongxiang Wang: Marshall School of Business, University of Southern California, Los Angeles, California 90089; email: Yongxiang.Wang@marshall.usc.edu. Rong Xu, School of Finance, Renmin University of China, 59 Zhongguancun St, Haidian, Beijing, China; email: xurongfsi@126.com. We would like to thank Harry DeAngelo, David Hirshleifer, Danielle Li, and Jerry Hoberg for useful comments. Wang and Xu thank the China National Science Foundation (Grant Number: 71472180) for financial support, and Wang and Shi thank the Australian Research Council (DP150102339) for financial support. Teng Zhong, Roy Ruoyu Ji, Jin Wang, Wenwei Yang, Hongyu Zhang, Tao Lu, and Eric Hardy provide very able RA work.

Economists have long considered the consequences of rent-seeking and the resultant allocative inefficiency for economic growth. Distortions in resource allocation may result from rent-seeking by firms or government officials (see, for example, Mauro, 1995), or talent across sectors (Murphy, Shleifer, and Vishny, 1991). In earlier work, emphasis has been on the misallocation of capital (human or physical) in production. But endogenous growth models, starting with Romer's (1986) landmark paper, emphasize knowledge production as the source of increasing returns required to generate growth in the long run. So rent-seeking and distortions in knowledge production – a topic that has seen relatively little discussion or empirical analysis – is of particular relevance to models of economic growth and development.

In this paper, we provide an empirical analysis that lies at the intersection of these two themes: rent-seeking and allocative inefficiency, and knowledge production. Our focus is on favoritism and the allocation of scientific talent in China. The setting is of particular relevance because of the increasingly prominent role that China plays in the global economy, and because of concerns over the long-term viability of Chinese economic growth (Zhu, 2012). The importance of scientific innovation for sustained growth is well-recognized by the Chinese government. As part of its strategy for economic development, the country has channeled over a trillion dollars into promoting scientific education and research over the past two decades (China Statistical Yearbooks of Science and Technology, 1996-2014). R&D expenditures have grown at more than 20 percent annually in recent years, and China is currently second only to the U.S. in R&D spending (Ni, 2015).

But press accounts have suggested that some of these expenditures have been misdirected as a result of favoritism and corruption. These stories implicate scholars and officials at the very highest levels. For example, in a widely-reported embezzlement case in 2014, Ning Li, a fellow of the Chinese Academy of Engineering (CAE) was among those convicted of misappropriating funds of 20 million RMB (3.17 million dollars)[1]. Corruption is thought to extend to the fellow selection process of the CAE and its more prestigious sister organization, the Chinese Academy of Sciences (CAS). Membership in the CAS confers considerable prestige – the title of CAS fellow is the highest official honor for Chinese scientists – as well as privileged access to research funds. These translate into direct material rewards as well, as CAS members are sought after (and well-compensated) by employers as a result.

There have been well-publicized examples of scientists attempting to gain CAS membership through bribery. In one notorious case, Shuguang Zhang was convicted of accepting bribes totalling 47 million RMB (about US$7.5 million) in his capacity as vice chief engineer of China's railway ministry, and using nearly half the proceeds to try to buy membership in the CAS. He came up one vote short in his second attempt in 2009, despite never having

[1] "Scientists caught in Chinese anti-corruption sweep," *Nature*, October 16, 2014)

published a peer reviewed journal article.[2]

Leading scientists have suggested that the problem runs much deeper than a few high-profile cases of outright corruption, arguing that the CAS/CAE selection system is opaque and dictated more by personal connections and lobbying than scientific achievement.The distorting effects of connections in Chinese science were described most forcefully by two prominent Chinese scientists, Yi Rao and Yigong Shi, in *Science* in 2010 (see Shi and Rao, 2010), where they suggest that "[a] significant proportion of researchers in China spend too much time on building connections and not enough time attending seminars, discussing science, doing research, or training students."[3] Further, Rao and Shi argue that once scientists attain positions of power and influence, "[s]ome become part of the problem: They use connections to judge grant applicants and undervalue scientific merit." They thus raise concerns of possible misallocation of effort by scientists (lobbying versus research) and also for misallocation of resources across scientists (from good researchers to effective lobbyists). Their narrative, furthermore, suggests that Chinese science may have settled into a "rent-seeking equilibrium" as described by Acemoglu (1995), in which the rent-seeking choices of today's scientists affect the rent-seeking incentives of future scientists.

In the empirical analysis that is the focus of our paper, we provide evidence of favoritism in the selection of candidates for membership in the CAS and CAE during the period 2001-2013, using a form of connections that plays a central role in Chinese society: hometown ties. We focus on hometown ties, or *laoxiang guangxi* in Chinese, because of their importance in the culture of favor-exchange (*guangxi*) in China, and because they are observable to us as researchers.[4] We measure connectedness in the fellow selection process based on whether the nominee's hometown is shared by a member of the standing committee in the department where the candidate is nominated for membership (the CAS and CAE, as we detail below, are organized by department on disciplinary lines, such as chemistry, mathematics, and so forth).

We show that, during 2001-2013, the probability that a nominee to the CAS or CAE is elected as a fellow is 36 percent higher if he is connected according to our measure. This result is highly robust. It survives the inclusion of year-times-department fixed effects, city fixed effects, undergraduate college fixed effects, and employer fixed effects, indicating that we are not picking up unobserved differences in scientist quality across city of origin, undergraduate

[2]Zhang received a death sentence after his record of bribe-taking was uncovered. The sentence was suspended for two years, and he remains in prison as of this writing. His case is detailed in "The True Cost of Becoming an Academician in China?" *ScienceInsider*, September 17, 2013.

[3]Both failed to get elected to the CAS in 2011.

[4]Beyond the central importance of such ties in Chinese society, this approach has some precedence in social science research. Siegel (2007), in particular, exploits regional ties in his analysis of favoritism by Korean government officials. A handful of studies in finance and economics use school ties as a measure of personal connections between companies and their investors (Cohen, Frazzini, and Malloy, 2008 and 2010), and among politicians in the U.S. Congress (Cohen and Malloy, 2014).

institution, or employer. Further, we find no effect from hometown ties to fellows who are not members of the nominee's department standing committee, nor any effect from hometown ties to members of standing committees from other departments. These 'placebo' results indicate that the higher success rate is quite specific to hometown ties to influential members of the candidate's own department. Finally, , we do not find a robust effect of connections to fellows via candidates' undergraduate institutions, which suggests that our results are unlikely to result from "soft" information on candidate quality, which would more plausibly be captured by shared academic rather than hometown connections.

We next disaggregate the role of hometown ties into the effect on each round in the two-stage fellow selection process. In the first stage, where evaluations are done by mail by a broader set of CAS/CAE members within each department (and where the main purpose is to filter out obviously unqualified candidates), we find no effect of hometown ties to the standing committee. Rather, first-stage selection is well-predicted by candidates' publication records, as proxied by their H-Indexes at the time of nomination. The hometown effect comes exclusively from the second stage, where final selection is conducted in an in-person meeting, a setting where personal lobbying and influence more plausibly play a role in voting behavior. (In this second stage, publication records are not predictive of success.)

If hometown-connected fellows face a lower threshold for election, two further predictions follow (a) the average quality of a connected nominee may be lower; and (b) conditional on the pool of nominees, the quality of elected fellows will be lower. Based on analyses of candidates' H-Indexes, we find no support for the former, i.e., nominee quality is unrelated to connections. However, we find strong support for the latter prediction: among nominees without hometown ties, H-Indexes are much higher among elected than non-elected candidates (i.e., there is positive selection on publication quality); for connected candidates, however, elected candidates' H-Indexes are no higher than those of non-elected ones (i..e., no positive selection).

Finally, in examining the selection process over time, we find that the higher probability of election enjoyed by hometown-connected nominees virtually disappears in 2007. That year, owing to outside pressure to increase transparency and fairness in fellow selection, the CAS and CAE changed the election rules to increase the fraction of 'yes' votes required for a nominee to be appointed fellow, from one half to two thirds, and began publishing online lists of nominees and those ultimately selected. We speculate that the change in election rules may have made it more difficult for influential fellows to secure enough votes to gain approval for their favored applicants. (It is impossible, however, to separate in the data the effect of changed election rules from the effect of publishing election results; nor can we separate these effects from that of public pressure to clean up the selection process, the latent force that drove both changes.)

Overall, our results indicate that favoritism led to negative selection into the CAS and CAE, organizations that, as we detail in the next section, wield immense power in China's scientific community. While measuring the impact on knowledge production is beyond the scope of our paper, if our findings are reflective of a broader culture of favor-exchange in Chinese science (as Shi and Rao suggest they are), then the consequences for scientific innovation (and hence economic progress) are potentially significant.

Our work relates most directly to a growing literature on the role of personal bias on resource allocation in scientific research. Li (2015), for example, studies the role of reviewer "relatedness" in the awarding of National Institutes of Health funding in the U.S. Li finds that applicants connected to a reviewer via citation history are more likely to receive funding. She finds evidence of both better information and bias, though the informational effect dominates. More in the spirit of our findings are Zinovyeva and Bagues (2015) and Durante et al (2014), which document links between personal ties and faculty promotion in Spain and Italy respectively. Zinovyeva and Bagues (2015) look at promotion to associate or full professor, and focus on strong professional ties – such as whether an evaluator was the candidate's Ph.D. advisor – whereas Durante et al study the role of familial connections in promotion decisions, based on surname similarity. 2

Considered more broadly, we contribute to our understanding of rent-seeking and resource misallocation, an enormous literature within economics. Much of this work emphasizes rent-seeking and the allocation of human capital across sectors or activities, as in Murphy, Shleifer, and Vishny (1991), Acemoglu (1995), and many others. We study favoritism in knowledge production, thus linking this body of work on rent-seeking to the topic of knowledge production and endogenous growth theory.

Empirically, work on resource misallocation in China (and elsewhere) has focused on misallocation across firms, a concern first given prominence in the economics literature by Young (2003). Hsieh and Klenow (2009), in particular, document large gaps in the marginal product of labor and capital in China versus the United States, and argue that reallocation of resources across Chinese firms could increase manufacturing total factor productivity (TFP) by 30 - 50 percent. Khandelwal, Schott, and Wei (2013) document misallocation of export licenses, showing that liberalization led to increased exports from more productive firms that had previously been shut out of the export market. (If export quotas were allocated based on efficiency, one would not expect this result.) Brandt et al (2012), like Hsieh and Klenow, study productivity in China's manufacturing sector. While confirming Hsieh and Klenow's results on the low *level* of TFP, they find substantial improvements in TFP during 1998 - 2007. (Our findings are similar in spirit to Brandt et al. (2012) in the sense that, while we also find evidence of allocation on the basis of personal ties rather than merit, this effect largely disappears in the latter part of our sample.) In contrast to this earlier literature, our study

provides more direct evidence on the allocation mechanism that is the source of distortions (hometown ties), and focuses on the entirely distinct domain of scientifc enterprise.

Our paper also relates to the very large literature on the role of culture in exchange. Much of this work focuses on trust in market transactions as in, for example, Guiso, Sapienza, and Zingales (2009) and Fisman, Paravisini, and Vig (2015). More closely related to our work, Parsons et al (2011) discuss how favoritism along ethnic lines can distort decisions, showing that Major League Baseball umpires make favorable calls toward players of their own ethnicity.

Finally, we contribute to the literature on the distortionary effects of in-group favoritism. Most directly related to our paper, there is a growing concern within China on the abuses of *guanxi* where, as we discuss in detail below, hometown ties play an important role (see Li (2010) for a general treatment and Lin et al (2015) for an economics-focused discussion in the context of recent anti-corruption reforms). The distortions of in-group favoritism is a global concern, however: Burgess et al (2015), for example, show in Kenya that districts that share the president's ethnicity receive twice the road building funds (and quadruple the length of paved roads) as unconnected districts. As with Burgess et al, our work provides a microeconomic underpinning to the macro-level association between ethnic fissures and corruption, as documented by, for example, Mauro (1995), Olken (2006), Easterly and Levine (2007), and Alesina and Zhuravskaya (2011).

1 Background

1.1 The Chinese Academies of Sciences and Engineering

The Chinese Academy of Sciences describes itself as "the linchpin of China's drive to explore and harness high technology and the natural sciences for the benefit of China and the world," and lays claim to "over 85 percent of China's large-scale science facilities" spread across over 1,000 CAS-affiliated sites throughout China. In addition to promoting science through its affiliated institutions, the CAS serves as an academic society, with membership in the CAS seen as the country's scientific highest accolade. As of 2014, the CAS had 711 members (including 274 emeritus members over the age of 80, who play no role in the selection of new members) spread across six divisions: Mathematics and Physics; Chemistry; Biological and Medical Sciences; Earth Sciences; Technological Sciences; and Information Technology Sciences (the last of these was carved out of Technological Sciences partway through our sample, in 2005). The Chinese Academy of Engineering (CAE), the CAS's sister organization, consists of nine departments (with 791 current members): Engineering Management; Chemical, Metallurgy and Material Engineering; Mechanical and Vehicle Engineering; Energy and Mining Engineering; Civil and Hydraulic Engineering; Light Industry and Environment Engineering; Information

and Electronic Engineering; Medicine and Health Faculty of Engineering; Agriculture; and Light Industry and Environment Engineering. The last two were created from a split of a single department in 2006.

It is hard to overstate the influence that fellows collectively have on the deployment of scientific resources in China. The CAS itself was given control of over US$400 million in research funds in the 2014 national budget for "strategic priority projects in areas ranging from neuroscience to studies of the Tibetan Plateau." Additionally, "megaproject" grants from the Ministry of Science and Technology, or MOST (which has a budget that is currently approaching US$10 billion) often require CAS or CAE fellow recommendations.[5] Further, MOST often draws on the CAS and CAE to fill its leadership ranks. For example, CAE fellow Ning Li ran the country's "National High-tech R&D Program," known as Project 863, until he was, as we mention in the introduction, arrested for embezzlement.

Beyond the honor of membership, fellows enjoy a range of material benefits. These range from chauffeur services to priority access at China's best hospitals (fellows have medical benefits comparable to vice-minister level government officials). Local provinces often augment the perquisites of CAS and CAE members in their efforts to lure them from the country's urban centers. For example, the CAE's website details the benefits of fellows residing in Hunan province, where a fellow's employer is required to provide a salary of at least 200,000 RMB, a starting research budget of at least one million RMB, and a car and driver. Specific employers can choose to further augment fellows' benefits. For example, Jinan University in Shangdong province explicitly states that the school will provide fellows with an annual salary of two million RMB, a moving allowance of one million RMB, and also a free home.[6]

Standing Committees within the CAS and CAE

Each department within the CAS and CAE has a standing committee, which plays a critical role in the fellow selection process. The committees are each comprised of 15-23 fellows, depending on the department's size. Standing committee members are nominated by fellows within each department, and one standing committee member is further elected as director of the department, along with 3 - 5 vice-directors (also from the standing committee's ranks). In its election rules, the CAS explicitly states that standing committees should maintain a

[5] The following link provides one such call for funding that requires three recommendations from CAS fellows (or researchers that hold other prestigious titles, such as Yangtze River scholar, the highest honor bestowed upon Chinese researchers): http://www.most.gov.cn/tztg/201008/t20100824_79062.htm. Downloaded on October 19, 2015.

[6] This information was taken from a job listing at Jinan University, posted on the school's official website. Unlike U.S. schools, many Chinese universities provide detailed compensation information when advertising job openings.

balance of membership based on subfields, sectors, and also regions.[7] In the CAS prior to 2008, members served two year terms for up to three terms; terms were then lengthened to four years, renewable only once. There is mandatory turnover: until 2008, at least a third of standing committee members had to be replaced every two years; starting in 2008 at least half of committee members needed to be replaced every four years.

Finally, there are academy-level committees (*Xubu Zhuxi Tuan* in Chinese) within both the CAS and CAE. These academy-level committees consists of directors of each department, the dean and vice-deans of the academy, and a few other fellows that are elected by a general vote at the biennial academy meeting. Elected committee members have four-year terms, which is nonrenewable in the CAS (renewable only once in the CAE). This committee is in charge of daily adminstration of the academy. The number of academy-level committee members in the CAS ranged from 27 to 38 during our sample period; for the CAE, the academy-level committee had between 31 and 37 members.

Selection of new CAS and CAE fellows

In the CAS, election of up to 60 new fellows across the six divisions takes place biennially in odd years, with the CAS-level standing committee deciding on the allocation of openings across departments. The CAE similarly elects up to 90 fellows across its nine departments, with the distribution at the discretion of the CAE standing committee.[8] Candidates may be nominated either by any existing fellow or via the candidate's employer. In the latter case, the nomination is then vetted by the ministry-level unit that oversees the employer, with the ministry deciding which nominations will be put forth amongst those under its administration. For example, Peking University is administered by the Ministry of Education, so Peking University may put forth nominations to the Ministry of Education which will assess these candidates and those from other universities, and then decide which university-affiliated candidates will receive formal nominations.[9]

Within each department, selection amongst these nominees is overseen by the standing committee. Selection proceeds in two main stages. First, each standing committee organizes several subgroups within its department based on academic expertise (e.g., organic and inorganic chemistry), with each subgroup including at least 15 fellows, to provide individual

[7] The CAS website provides details of the standing committee election process: http://history.casad.cas.cn/document.action?docid=11998 (downloaded September 14, 2015)

[8] See the CAS bylaws for details on the current selection process http://english.casad.cas.cn/Ab/Re/ (downloaded October 31, 2015)

[9] According to CAS/CAE bylaws, the ministry-level units that may nominate candidates include a number of central government ministries in Beijing (i.e., Ministry of Education, MInistry of Agricutlure, Ministry of Finance, and so forth), all provincial governments (including Beijing, Tianjin, Shanghai, and Chongqing), the China Science Association (a ministry-level unit), and the four departments of The People's Liberation Army (General Staff Department, PLA General Political Department, PLA General Logistics Department, and PLA General Armaments Department).

written evaluations of applications along with a yes-no vote. Every member of the department receives these subgroup assessments (along with the final votes of each subgroup member), and is then required to provide a yes-no vote on every candidate in the entire department. This department-wide vote is used to weed out about 40 percent of the initial pool of applicants. We refer to this winnowing as the first selection stage.

The second stage begins with an evaluation of the remaining applicants by a group of three fellows selected by the department standing committee (and potentially including standing committee members themselves), who then present their evaluations to the entire department. Voting then proceeds in two steps. First, all participating fellows vote on the set of candidates that made it through the first stage, and based on these votes a short list of "formal candidates" for fellows is generated. The number of formal candidates is equal to 1.2 times the number of available slots in each department (the multiple was 1.4 prior to 2008). Finally, in the second step of this stage, there is a new round of voting by all participating fellows. Candidates are ranked based on the number of yes votes received, with the highest ranked candidates selected as fellows as long as they receive yeses from at least two-thirds of votes cast (prior to 2006, candidates needed to receive a yes from just one half of voting fellows).

This process gives standing committee members considerable sway in the selection of fellows, potentially across each stage of the process. In the first stage, they assign candidates' applications to fellows within their departments for initial review. For the second stage, standing committee members organize the three-person group which evaluates each remaining candidate (likely including at least some committee members themselves), and have a chance for exercising social pressure in the final in-person vote. Finally, we observe that while non-standing committee members may not be present for the vote, standing committee members, as department leaders, are required to be present, and their comments on each candidate at the outset of the meeting can set the tone for the subsequent discussion.[10]

The election rules at the CAE are very similar to those of the CAS, subject to minor differences. Most notably, candidates who make it past the first stage are required to do a presentation (and answer questions from current fellows) prior to the second stage vote.

1.2 The importance of hometown ties in China

Hometown ties, or *laoxiang guanxi* in Chinese, play a central role in *guanxi,* the culture of favor exchange in Chinese society. As expressed by anthropolist Leo Douw in an introduction to a book-length treatment of the topic,"[t]he cultivation of hometown ties is part and parcel

[10]Once each department has selected its fellows, final academy-wide approval is required, but this step is largely a formality. Each department sends its fellow list to the CAS-level standing committee for procedural approval. After 2014 (our data end in 2013), the election rules shifted somewhat, and approval of all candidates required a CAS-level vote, though this too was seen as largely pro forma.

of the Chinese culture of establishing *guanxi*, or relationships of mutual obligation between individuals, and is therefore also an inherent part of the social structure in which doing business in China is embedded at present. Moreover, ethnic Chinese communities abroad have usually preserved a distinctly Chinese cultural identity which is centered on the sharing of roots in the hometown." As Chen and Chen (2004) observe, hometown ties are among the most common and distinctive bases for *guanxi* to build upon.[11]

There is a literature too vast to survey here that examines the origins of *laoxiang guanxi*, and also documents the many roles that it plays in contemporary Chinese society. Social organizations based on place of origin are very common amongst immigrant groups, and are used to facilitate communication, strengthen within-group networks, enlarge the group's political power in the new location, and also to form coalitions to better compete in commercial enterprises. There are often formal organizations built around *laoxiang guanxi*, typically called *Tong xiang hui* or *Lao xiang hui* (hometown associations) or *Huiguan* (guildhouses). These formal associations are common both among migrant communities within China and also amongst the global Chinese diaspora.[12]

These connections have led to favor exchange that has been explicitly censured by government officials. In early 2015, a director at the Central Commission for Discipline Inspection, China's highest anti-corruption authority, expressed his concerns on the culture of favor-exchange that has arisen around hometown networks, worrying that government officials maintain their hometown ties solely for the purpose of building profitable connections to businesses or gaining promotion.[13] In October 2015, the Communist Party of China acted on these concerns by banning hometown associations altogether, under the rationale that they served to facilitate corruption amongst government officials, and between businesses and government officials.

At a highly anecdotal level, hometown ties have been implicated in corruption of the CAS/CAE selection process that is our focus. In October 2013, Sciencenet, a publication cosposored by the CAS and CAE, reported on the case of Mingxian Chen, who in 2011 was the vice-chief officer of Hunan province's Transportation Department.[14] Chen was nominated that year for the civil engineering department of the CAE by a standing committee member from his home province of Hunan after trying (and failing) to secure a nomination from a fellow from his hometown of Changde City. His nomination failed in the second stage after

[11] Hometown networks play a prominent role among migrant groups from other countries as well. See Munshi (2003), for example, on the role of hometown networks among Mexican immigrants in the U.S.

[12] See, for example, Ho (1966) for a classic account of the social and political role of *Huiguan* in China; Dou (1946) for details on *Tongxianghui* within China; and Moll-Murata (2008) for a discussion of Chinese guilds going back to the seventeenth century. For discussions of hometown-based associations amongst the Chinese diaspora see, for example, Freeman (1960); Crissman (1967), and Kerri 1976).

[13] See http://www.hebgcdy.com/2015/0123/104692.html accessed on Oct 6, 2015

[14] See http://news.sciencenet.cn/htmlnews/2013/10/283957.shtm accessed October 9, 2015.

it came to light that some of his research contributions were fabricated or written by others. (Chen was arrested for corruption in 2012.) The experiences of Chen might thus be seen as one indication both of efforts to use hometown (or even home province) ties gain entry to the CAS/CAE, but also the greater scrutiny that candidates (and their connections) have received in more recent years.

While it is beyond the scope of our empirical exercise to document the direct welfare effects of hometown favoritism, we conclude this section by reiterating that the consequences are potentially quite substantial for knowledge production in China, and need not be limited to the selection of lower-quality fellows (one of our findings below). The CAS wields enormous influence over scientific funding and resource allocation more broadly. Fellows willing and able to exploit connections in obtaining CAS/CAE membership might plausibly continue to rely on hometown ties as a conduit for further favor exchange.

2 Data

Candidate information for both the CAS and CAE were obtained from the organizations' official websites (www.cas.cn and www.cae.cn) and the CAS's official publication *CAS Bulletin*, where these data have been published since 2001. The CAS and CAE sites provide information on a candidate's passage through both the first round of initial screening as well as final selection. (There is no information on whether a candidate was included in the short list that was considered in the final in-person vote, consisting of 1.2 times (pre-2008) or 1.4 times (post-2008) the number of available positions.) There are two exceptions: The CAS has not posted the listed of candidates who passed through the first stage for 2001 and 2013, while the CAE has not posted this information for 2001. We filed requests for this information via China's freedom of information laws ("Regulation of the People's Republic of China on the Disclosure of Government Information," in effect since since May 1, 2008). The CAS has not complied with our request, responding that the information "is not a required disclosure under the government's information law."[15] The CAE sent us the data from 2001.

The nominee lists we obtained from these sources are used to construct our two main outcome variables. $Elected_{yi}$ is an indicator variable that denotes whether candidate i in year y was elected a fellow. We also generate the indicator variable $FirstStage_{yi}$ which denotes whether candidate i passed through the first stage of the selection process in year y. (Recall that we do not have information on the first stage of selection for the CAS for 2001 and 2013.) It is not uncommon for applicants to fail in their first bid at membership in the CAS and CAE, so a single candidate i may appear in multiple years (by definition, with

[15] Translations of this correspondence are available from the authors.

Elected = 0 in all but the final year). Of our sample of candidates, 1553 (54 percent) are nominated only once, 777 (27 percent) are nominated twice, and 546 receive nominations three or more times. Conditional on receiving a nomination after an initial failure, a candidate's success rate is much higher. For example, the success rate is 9.5 percent for all candidates in their first attempt, versus 17.7 percent for candidates who are re-nominated on their second attempt.

Nominees to the CAS and CAE – even the unsuccessful ones – are generally well-known individuals, often members of the scientific and social elite. We were thus able to obtain personal and professional information on nearly all CAS nominees and a majority of CAE nominees. For all nominees, information was obtained through a combination of employer websites, listings on *Baike* (China's Wikipedia, which is a subsidiary of *Baidu*, China's Google). These sources were generally sufficient to obtain a candidate's birth year, gender, municipality of birth (including the rural area within the jurisidiction), and educational background. These sources were supplemented by, as needed, name searches via *Baidu* and also author listings in *ckni.net*, the Chinese version of JSTOR, as some Chinese journals require that authors provide age and city of birth. For elected fellows, the process was facilitated by the short biographies posted on the CAS and CAE official websites. We were unable to find city of birth for 555 candidates out of a total of 3334. Of these, 54 candidates (5.2 percent of the total of 1045) were CAS candidates, while 501 (21.8 percent of the total of 2289) were CAE candidates. These candidates are necessarily excluded from our analysis.[16] Finally, the CAS official website provides a listing of all standing committee members for each department for the period of 2001-2013. While the CAE official website does not provide this information, standing committee lists are provided in hard copies of CAE yearbooks from 2001-2013.

By combining city of birth information on both fellows and nominees with department standing committee listings, we generate the candidate-year level variable $CommitteeTie_{yi}$, denoting that candidate i in year y was born in the same city as at least one standing committee member in his department. We similarly generate $NonCommitteeTie_{yi}$, which denotes that a candidate was born in the same city as at least one fellow in his department but *not* on the standing committee. This variable captures, for example, the extent to which a particular city tends to produce high-quality chemists or mathematicians. We generate a further "placebo" measure of hometown ties that captures whether a nominee is connected to department committee members in departments other than his own, $CommitteeTie_Placebo_{yi}$.

Based on candidates' educational backgrounds, we similarly generate variables that indi-

[16] There are no significant differences in age or H-Index between CAS/CAE candidates where we were able to find birthplace information versus candidates where we could not. The average age is near-identical for the two groups: 58.4 for those with hometown information, 58.4 for those without. There is similarly no difference in average H-index (8.8 versus 8.4 for those with and without hometown information, p-value of the log difference between the two of of 0.70).

cate that a candidate attended the same undergraduate institution as a standing committee member ($Committee_CollegeTie_{yi}$) or that a candidate attended the same undergraduate institution as a fellow not on the standing committee ($NonCommittee_CollegeTie_{yi}$). In a similar vein, we generate $Committee_EmployerTie_{yi}$ to denote whether a candidate is employed at the same institution as a standing committee member at the time of nomination, and $NonCommittee_EmployerTie_{yi}$ to denote that a candidate is employed at the same institution as a fellow not on the standing committee.[17]

Throughout our analysis, we wish to control for academic output, which we proxy for using each candidate's H-index at the end of the relevant election year, obtained from Web of Science (Core Collection). While there is no sufficient statistic for observable candidate quality, the H-Index is an accepted measure that captures both quantity of output as well as citation impact (Hirsch (2006). Our H-Index is calculated for 2014, but only includes work published by the end of the year of nomination. This allows us to incorporate a forward-looking view of publications of relatively recent vintage.[18] In many cases, there were multiple search results due to candidates' use of initials rather than full names, or common names. In these cases, we matched also based on author field and affiliation. We transform the resulting variable, $HIndex_{yi}$ using $log(1+x)$ because of the long right tail of the distribution, and because 27.2 percent of candidate-year observations had an H-Index of zero, indicating that they failed to produce any publications, or published only in Chinese journals not included in H-Index calculations. As a further measure of academic credentials, we include $HasPhD_{yi}$, an indicator variable dentoing that a candidate holds a Ph.D. or equivalent degree. (We caution that the lack of a doctoral degree is not in itself an indication of inadequate qualifications. For example, the 2015 Nobel laureate in medicine, Youyou Tu, did not hold a doctoral degree. Doctoral degrees are, as we discuss below, far more common among younger nominees.)

We include several further controls to account for other forms of status and connections. $Dean_{yi}$ indicates that a nominee holds an administrative rank of dean or higher at his academic institution, while $PoliticalTie_{yi}$ captures whether the nominee is (or was) a *Tingju* level (or above) government official, where a *Tingju* level official holds the same rank as a city mayor. In our context, politically influential candidates are typically former government officials (including some of very high rank; for example, the former Minister of Railway, Zhihuan Fu, was a candidate in 2001 and elected as a fellow in that year) as well as high-ranking members of the military.

[17] The CAS itself operates 84 largely autonomous institutes spread throughout China with each one typically specializing in a particular scientific subfield. We treat CAS-affiliated nominees in different municipalities as having separate employers for the purposes of this variable's construction.

[18] In this, we follow the innovations literature, which typically allows for several years' lag in measuring citations. See, for example, Aghion et al. (2013). Additionally, our forward-looking measure of publication impact at the time of election incorporates any positive treatment effect that CAS/CAE membership has on citations, but this should be symmetric in its impact on connected and unconnected fellows.

In Table 1A we list summary statistics for the full sample, while in Tables 1B and 1C we provide summaries disaggregated by $CommitteeTie$ and $NonCommitteeTie$ respectively. Note that the latter two groups are not mutually exclusive – a candidate may have both committee and non-committee ties, and this is in fact not uncommon in our data.

There are a few patterns worth highlighting in the comparison of candidates with and without committee hometown ties. First, in the raw data there is a gap of 6 percentage points in the fraction of nominees that are elected fellows between $CommitteeTie = 1$ candidates and $CommitteeTie = 0$ candidates (19.7 percent versus 13.7 percent, difference significant at the 1 percent level). However, we observe no difference in the fraction of candidates who make it past the first stage of selection. In fact, $CommitteeTie = 0$ candidates enjoy a slightly higher success rate in the first stage of screening (40.6 percent, versus 38.8 percent for $CommitteeTie = 1$ applicants), though this difference is not statistically distinguishable from zero. It thus follows that there is a very large difference in $Elected|FirstStage = 1$, the fraction of candidates elected conditional on making it past the first stage. Its value is 17.5 percentage points higher for $CommitteeTie = 1$ candidates (49.7, versus 32.2 for $CommitteTie = 0$ candidates). The difference in success rates based on non-committee ties are much more modest and, as we will see in our regression results in the following section, do not survive the inclusion of basic controls.

The second point to note in Tables 1B and 1C is that there are a number of other sharp differences between hometown connected and unconnected candidates. These all stem from two main differences: First, there are cities that tend to produce a large number of scientists, who also tend to go to elite academic institutions. Hence, in particular, in both tables we observe a large difference in means for school ties, as evident in the last two rows of each table. This will make it particularly important to ensure that our results are robust to city-of-origin fixed effects and also to consider the effects of "placebo" measures of connections that capture the scientific strength or prevalence of scientists from particular locales.

Additionally, there is a significant age difference in both tables between conncted and unconnected candidates. The difference between candidates with hometown committee ties and those without is 1.5 years; the age gap is even wider for non-committee hometown ties, where the mean difference is 2.3 years and highly significant. The difference in the fraction of candidates with Ph.D.s is a direct result of this age gap. Doctoral degrees were uncommon among Chinese researchers until relatively recently. For example, over 70 percent of candidates under the age of 60 hold doctoral degrees (almost 85 percent of applicants under 50), while the rate is below 20 percent for candidates 60 and over. Once we control for age, the difference in fraction with doctoral degrees in Tables 1B and 1C disappears.

This still leaves the question of why there exists an age difference based on connections in the first place. While highly speculative, we can offer one interpretation for these age gaps.

Recall that candidates may be nominated by employers or current fellows. While we do not observe the source of a candidate's nomination, an employer would plausibly be less inclined to nominate a candidate as he approaches retirement, leaving only current fellows as potential nominees. The higher age of nominees might thus be a manifestation of favor-seeking among related fellows in the nomination process, which is necessarily more prevelant among older cohorts. This is roughly consistent with the age profile of connected nominees, as illustrated in Appendix Figure A1, which shows the fraction of each 5 year cohort of nominees that are connected to current fellows. The fraction dips briefly, then rises steadily to level out at the 65-69 cohort. The mandatory retirement age is 60.

We will control for $log(Age)$ in our main specifications below, and also show results with a full set of age cohort fixed effects, which generates virtually identical point estimates and standard errors.

3 Empirical Results

We begin in Figure 1 by showing how the gap in election rates by hometown connection status varies across time. We divide the sample into three non-overlapping groups: $CommitteeTie = 1$ candidates; candidates lacking hometown ties to their department standing committee but with ties to non-committee department fellows ($NonCommitteeTie = 1$ and $CommitteeTie = 0$); and candidates with no hometown ties to department members at all (i.e., both hometown tie variables are zero). Two noteworthy patterns emerge. First, we observe virtually no difference in election rates between the two groups where $CommitteeTie = 0$, suggesting that there is no benefit from connections to non-committee members, and that this lack of benefit is consistent over time. Second, $CommitteeTie = 1$ candidates have substantially higher election probabilities than both 'control' groups only prior to the 2007 election: In the earlier part of the sample, the success rate of candidates with standing committee hometown ties is about two-thirds higher than that of candidates without such ties. In 2007 the election rates of all groups falls, but the drop is far greater for $CommitteeTie = 1$ candidates. By 2009, the groups have fully converged, driven by a decline in the success rates of candidates with hometown ties to the standing committee, rather than an improvement in the success rates of candidates without such connections.

The are several possible explanations for the sudden drop in the success rates of candidates with committee hometown ties (and the decline in the average success rate of candidates overall). Given the sharp convergence in 2007, we look for potentially significant changes around that date. We speculate that an important factor may have been the change in the fraction of yes votes required for election in the second stage of the selection process, from

one half to two thirds. This could account both for the general decline in election rates, and also the disproportionate impact on hometown-connected candidates, as it plausibly made it more difficult for influential fellows to secure enough votes to gain approval for their favored applicants (assuming that the cost of securing an incremental vote increases more rapidly for candidates elected on the basis of connections rather than merit). There were other concurrent changes that might have also affected candidate selection. Starting with the 2007 election, the candidate lists for both the CAS and CAE have been published in two national newspapers, the *People's Daily* (the highest-circulation paper in China) and the *Guangming Daily*. Prior to 2007, candidate lists were available for the CAS via its internal newletter, the *CAS Bulletin*, and on the CAE's own website. Both the increased publicity and changes in electoral rules were likely in response to public criticism of the CAS/CAE fellow selection process. We cannot distinguish in the data between the effects of publishing candidate lists, changed electoral rules, and public demands for greater transparency in the selection process.

We now turn to a regression analysis of candidate selection in Table 2, showing successively more demanding specifications. Our main specification takes the form:

$$Elected_{yi} = \alpha_{dy} + \beta_1 * CommitteeTie_{yi} + \beta_2 * NonCommitteeTie_{yi} + Controls_{yi} + \epsilon_{yi} \quad (1)$$

where α_{dy} is a set of department-by-year fixed effects (7 years by 15 departments for both CAS and CAE, or a total of 105 fixed effects) and ϵ_{yi} is an error term that allows for clustering by candidate, since a single individual may apply multiple times.

In column 1, we include a number of covariates: $log(1 + Hindex)$ and $HasPhD$ to proxy for candidate quality; controls for political and academic stature via *Dean* and *PoliticalTie*; $Committee_CollegeTie$ and $NonCommittee_CollegeTie$ to capture whether a candidate went to the same undergraduate institution as fellows in his department; and $log(Age)$. We additionally include year fixed effects and department fixed effects (a total of 15 for the six CAS departments and nine CAE departments) to account for broad time trends and differential concentrations of hometown ties and success rates across departments. The coefficient on $CommitteeTie$ is 0.049, significant at the 5 percent level (p value of 0.012). Recall that the mean success rate of $CommitteeTie = 0$ candidates is 0.137, so that a hometown tie increases the probability of becoming a fellow by about 36 percent. This is sizeable when benchmarked against the impact of improved candidate quality as captured by the H-Index. The coefficient on $log(1 + HIndex)$ is 0.031 (significant at the 1 percent level), implies that a one standard deviation (1.256) increase of $log(1 + Hindex)$ raises the probability of being elected by 3.9 percentage points. The value of connections is thus comparable to a 1.26 standard deviation improvement in reseasrch productivity. (Though as we will see in Tables 3 and 4, the roles of

connections and research productivity manifest themselves in different stages of the selection process.) The coefficient on *NoncommitteeTie* is small in magnitude (0.005) and insignificant. We may reject the equality of coefficients on *CommitteeTie* and *NoncommitteeTie* at the 10 percent level.

In column 2 we add year-by-department fixed effects, and find that *CommitteeTie*'s coefficient is largely unchanged. To show that our results are robust to the inclusion/exclusion of covariates, in column 3 we repeat this specification including only year-by-department fixed effects and the two hometown tie variables, *CommitteeTie* and *NonCommitteeTie*. The point estimate on *CommitteeTie* is again virtually unchanged.

In column 4 we provide our most rigorous specification which includes hometown fixed effects for each of the 524 municipalities (including county cities) with at least one candidate during our sample, as well as year-by-department fixed effects. This captures any time-invariant differences in city of origin that might influence both the chances of serving on a department standing committee, and also success as a CAS/CAE candidate. The coefficient on *CommitteeTie* remains largely unchanged, though in this saturated specification the standard error also increases so the coefficient is significant only at the 10 percent level (p-value of 0.052). In column 5 we include fixed effects for candidates' undergraduate institutions. This leads to a modest increase in the coefficient on *CommitteeTie*, to 0.064, while the standard error is unchanged relative to our main specification in column 3. Finally, in columns 6 and 7 we separate the sample into CAS and CAE applicants; we find no difference between the two groups in the effect of committee hometown ties. Our measures of college ties mostly have positive coefficients, but they are generally insignificant, and turn negative in the specification with college fixed effects.

Beyond the inclusion of *NonCommitteeTie* as a control, we present a further sort of 'falsification test' in Appendix Table A1, which considers the effect of having hometown ties to a standing committee member *not* in the applicant's department (*CommitteeTie_Placebo*). These non-department standing committee ties have no effect on a candidate's election prospects, as one would expect given that election is done at the department level. We also show that our findings are robust to the inclusion of a full set of age cohort fixed effects in Appendix Table A2. The point estimates on β_1 are near-identical to those in Table 2, across all specifications, and the standard errors unchanged. In Appendix Table A3 we include *Committee_EmployerTie* and *NonCommittee_EmployerTie*, which capture nominees' ties to fellows through their employer at the time of nomination. We argue that these professional ties would more plausibly serve as a channel for soft information than hometown ties. The inclusion of employer-based connections does not change the size or significance of the effect of hometown ties. Further, the effect of employer ties to the standing committee is less robust and, in particular, does not survive the inclusion of employer fixed effects (column

5).

Finally, in Appendix Table A4 we allow for the effect of committee and non-committee ties to change in 2007, to examine whether the patterns in Figure 1 hold up in a regression framework. In column 1 of Table A4, we show that the success rate of connections is higher in the pre-2007 data – the point estimate on the direct effect of $CommitteeTie$ in column 1 is 0.088 – while its interaction with a post-2007 indicator variable, $I(year \geq 2007)$, is of near-equal but opposite sign (-0.091) and significant at the 5 percent level. (The remaining columns of Table A4 show how the results in Tables 3 - 5 change pre versus post-2007.)

We next separate the overall impact of hometown connections on selection into the first and second stages of the process. Our sample is smaller for these analyses relative to those presented in Table 2, because we were unable to obtain results from the first stage of selection for CAS candidates in 2001 and 2013. (In Appendix Table A5, we repeat the analyses of Table 2 to show that our main result on hometown committee ties is near-identical for this smaller sample where first stage data are available.) We present the results in Tables 3 and 4 for the first and second stages respectively. Focusing on the first stage, there are two interesting patterns that emerge in Table 3. First, the link between observable candidate quality (as capture by the H-Index) and progressing past the first stage is extremely strong. The coefficient on $log(1+HIndex)$ is 0.068 (p-value < 0.001), more than twice as large as the coefficient reported for overall selection in Table 2. There is similarly a positive association between $HasPhD$, plausibly another observable measure of candidate quality, and first stage success.

There is, however, no correlation between hometown ties and candidate success at the first stage. The point estimate on $CommitteeTie$ is close to zero in all specifications, and never significant. Based on $CommitteeTie$'s coefficient (-0.019) and its standard error (0.028) in column 2, our favored specification, we can rule out at the 95 percent level the existence of a positive effect of $CommitteeTie$ of greater than 3.6 percentage points (-0.019 + 0.028*1.96). In the first four columns of Table 3, we find a positive effect of college ties on first stage success, particularly for $NonCommittee_CollegeTies$. But this turns out to be a result of differences in college quality - once we add college fixed effects in column 5, the result disappears.

Given the lack of any first stage effect in Table 3, our results in Table 4 follow naturally: there is a very large effect of $CommitteeTie$ on second stage success across all specifications: a committee hometown tie is associated with a 15.3 percentage point increase in the probability of becoming a CAS/CAE fellow, conditional on making it through the first stage screening. Interestingly, $log(1 + HIndex)$ is only weakly correlated with sucess in the second stage. Its coefficient in column 2, for example, is about of quarter of the analogous estimate from Table 3.

Overall, our results in Tables 3 and 4 are consistent with a more prominent role for individual lobbying in the second stage which, as we observed in Section 1, takes place in a closed-door meeting. This stands in contrast to our finding that committee hometown ties are irrelevant in the first stage, where individual written evaluations dictate the outcome (though standing committee members choose the set of first stage evaluators, so this non-result is not obvious ex ante).

If hometown-connected fellows face a lower threshold for election, two further predictions follow: (a) the average quality of a connected nominee will be lower; and (b) the quality of elected candidates (conditional on the pool of nominees) will be lower for connected candidates. In Appendix Table A6, we examine the first of these predictions based on specifications paralleling our earlier results, with a candidate's $log(1 + Hindex)$ as the dependent variable. Across all specification, we find no effect of $CommitteeTie$ on $log(1 + HIndex)$. Thus, we find no relationship between hometown ties and observable candidate quality among the pool of nominees.

In Figure 2 and Table 5, we explore whether the quality of elected candidates (relative to the pool of nominees) differs for connected versus unconnected applicants. In Figures 2A - 2C, we provide a series of bar graphs to illustrate this relationship for overall selection, and disaggregated into success at the two stages of the process. Figure 2A shows that, for unconnected nominees, the median H-Index among elected candidates is twice as high as for non-elected candidates. By contrast, for connected candidates the two are virtually identical (and near-identical to the median H-Index for non-elected unconnected nominees). Figure 2B shows that, for the first stage, elected candidates have higher median H-Indexes for both connected and unconnected nominees, but the difference is greater for unconnected candidates. In the second stage (Table 2C), among connected candidates the median H-Index is actually *lower* among those elected as fellows (there is no difference among unconnected candidates).

We present regression results on overall selection in Table 5, augmenting our specification in Table 2 with the interaction term $CommitteeTie*log(1+HIndex)$. In column 1, we include department-times-year fixed effects, and our full set of controls from Table 2 (we suppress these coeffficients to conserve space). The direct effect of $log(1 + HIndex)$, capturing the association for unconnected candidates, is positive and significant at the 1 percent level. The coefficient on the interaction term, $CommitteeTie * log(1 + HIndex)$, is virtually identical in magnitude and opposite in sign (significant at the 5 percent level). In column 2 we add $NonCommitteeTie*log(1+Hindex)$ as a covariate, and in column 3 we add the interactions of $Committee_CollegeTie$ and $NonCommittee_CollegeTie$ with $log(1 + HIndex)$. Our main results are unchanged (and indeed in columns 2 and 3 even imply negative selection of connected candidates, as the sum of the coefficients on $log(1+HIndex)$ and $CommitteeTie*$

$log(1 + HIndex)$ is negative, though not significant). We include in columns 4 - 6 candidate hometown fixed effects; our results are largely unchanged.

Overall, Table 5 shows that the patterns in the raw data presented in Figure 2A – where we observed a positive association between H-Index and election for non-connected candidates but no such association for connected ones – are robust to a range of controls and specifications.[19] (In the final column of Appendix Table A4 we examine whether the interaction of $CommitteeTie$ and $log(1 + HIndex)$ loses significance after 2007. The coefficient on the triple interaction $CommitteeTie * log(1 + HIndex) * I(Year \geqslant 2007)$ is insignificant, however, and very close to zero. This non-result should be interpreted with caution, as our data are likely not up to the task of detecting this type of third-order effect. This is captured by the large standard error of 0.032 on this term.)

One alternative to our "favoritism" explanation - which fits many of the patterns we have described thus far - is that hometown ties allow for a better transfer of "soft" information on candidate quality (or potential to exploit the benefits of CAS/CAE membership to do better work). There are a few reasons we think this unlikely.[20] First, it is not clear what type of soft information on scientific qualifications would be resolved by a shared hometown, that is left uncaptured by (a) the expert assessments of department members in their stage one and stage two evaluations; (b) college and employer ties that are based more on academic links. One might be concerned that expert evaluation might not fully incorporate the quality of not yet published work. This is largely implausible, and not just because hometown ties are an unlikely conduit for the transmission of information on not yet published research: CAS and CAE appointments come later in individuals' careers (the median applicant is 59 years old), when sudden shifts in research output are less frequent.[21]

Overall, while we cannot rule out the possibility that hometown ties serve as a source of soft information, we argue that in practice it is very unlikely to explain our results.

[19] This indicates no complementarity between connections and ability. While this does not speak directly to the issue of soft information, it suggests – in the context of the model developed by Jia et al (2015) on promotion within the Chinese bureaucracy – that connections are more likely to be associated with loyalty than learning about a candidate's ability.

[20] On an anecdotal level, we would argue that if soft information played such a prominent role, it is extremely surprising that China's first and only science Nobel Laureate, Youyou Tu, was rejected in the first stage of selection in 2001, the year she was nominated for the CAE. While she had a weak publishing record, her Nobel prizewinning discovery on malaria treatment was made in the 1970s, and recognized as pioneering research even then.

[21] We examined whether candidates' H-Indexes in 2014 were correlated with hometown ties, similar in spirit to Li's (2015) approach of looking at future citations on NIH-funded work, depending on whether the candidate was connected to an award committee member. We find no effect. But this is a weak test, given that CAS/CAE fellowships are late career appointments.

4 Conclusion

In this paper, we study the fellow selection process for China's Academies of Sciences and Engineering. Nominees with hometown ties to department standing committee members were 36 percent more likely to be selected as CAS/CAE fellows, resulting entirely from higher success rates in the second (in-person) phase of the selection process. First and foremost, we view our work as providing some data and analysis to support the concerns that leading scientists have long expressed over favoritism and corruption in the Chinese scientific establishment.

The consequences are potentiallly significant, given the importance of knowledge production in the process of economic growth. And in the spirit of Acemoglu (1995), it raises concerns that Chinese science could settle into a rent-seeking equilibrium, as scientists advance through connection-building rather than innovation, and later disburse funding and other resources on the basis of connections, leading to further distortions in the incentives of future scientists.

We conclude, however, on a more optimistic note: the fact that the "hometown advantage" in fellow selection largely disappears in 2007 suggests that greater scrutiny and amended election rules may have been effective in curbing at least this form of favoritism, although it is possible that other channels of favoritism not observable to us are still present.

References

[1] Acemoglu, Daron. "Reward structures and the allocation of talent." *European Economic Review* 39.1 (1995): 17-33.

[2] Aghion, Philippe, John Van Reenen, and Luigi Zingales. 2013. "Innovation and institutional ownership." *American Economic Review*, 103(1): 277-304.

[3] Alesina, Alberto, and Ekaterina Zhuravskaya, 2011, "Segregation and the Quality of Government in a Cross Section of Countries." *The American Economic Review*: 1872-1911.

[4] Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, 2012. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing." *Journal of Development Economics* 97.2: 339-351.

[5] Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., & Padró i Miquel, G., 2015. "The value of democracy: evidence from road building in Kenya." *American Economic Review*.105(6): 1817-1851

[6] Cohen, Lauren, and Christopher J. Malloy. 2014. "Friends in High Places." *American Economic Journal: Economic Policy*, 6(3): 63-91.

[7] Cohen Lauren, Andrea Frazzini, and Christopher Malloy, 2008. "The Small World of Investing: Board Connections and Mutual Fund Returns," *Journal of Political Economy*, vol. 116(5), pages 951-979, October

[8] Cohen Lauren, Andrea Frazzini, and Christopher Malloy, 2010. "Sell-Side School Ties," *Journal of Finance*, vol. 65(4), pages 1409-1437, 08

[9] Crissman, Lawrence W. 1967, "The Segmentary Structure of Urban Overseas Chiense Communities", *Man*, New Series, Vol.2, No.2: 185-204

[10] Dou Jiliang, 1946, Tongxiang Zuzhi Zhi Yanjiu, [Study on the hometown associations], Taibei: Taiwan Zhengzhong Publishing House.

[11] Durante, Ruben, Giovanna Labartino, and Roberto Perotti. 2014. "Academic Dynasties: Decentralization and Familism in the Italian Academia." Unpublished manuscript.

[12] Easterly, William, and Ross Levine, 1997. "Africa's growth tragedy: policies and ethnic divisions." *The Quarterly Journal of Economics*: 1203-1250.

[13] Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2012. "Cultural proximity and loan outcomes." No. w18096. National Bureau of Economic Research.

[14] Freedman, Maurice, 1960, "Immigrants and Associations: Chinese in Nineteenth-Century Singapore". *Comparative Studies in Society and History*, 3(1): 25-48

[15] Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2009. "Cultural Biases in Economic Exchange?." *The Quarterly Journal of Economics* 124.3: 1095-1131.

[16] Hirsch, Jorge E, 2005. "An index to quantify an individual's scientific research output." *Proceedings of the National academy of Sciences of the United States of America* 102.46: 16569-16572.

[17] Ho, Ping-ti., 1966. Zhongguo Huiguan Shi Lun, [Study on the history of Chinese guildhouses]. Taibei: Taiwan Xuesheng Publishing House.

[18] Hsieh, Chang-Tai, and Peter J. Klenow, 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124.4: 1403-1448.

[19] Jia, Ruixue, Masayuki Kudamatsu, and David Seim. "Political selection in China: The complementary roles of connections and performance." *Journal of the European Economic Association* (2015).

[20] Kerri, James Nwannukwu, 1976, "Studying Voluntary Associations as Adaptive Mechnisms: A Review of Anthropological Perspectives". *Current Anthropology*, 17(1): 23-47.

[21] Khandelwal, Amit K., Peter K. Schott, and Shang-Jin Wei, 2013. "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters." *American Economic Review* 103.6: 2169-2195.

[22] Li, Danielle. 2015. "Information, Bias, and Efficiency in Expert Evaluation: Evidence from the NIH." Unpublished manuscript.

[23] Li, Ling, 2011. "Performing Bribery in China: guanxi-practice, corruption with a human face." *Journal of Contemporary China* 20.68: 1-20.

[24] Lockwood, Benjamin, Charles Nathanson, and E. Glen Weyl, 2014, "Taxation and the Allocation of Talent." Available at *SSRN* 1324424.

[25] Mauro, Paolo.1995, "Corruption and growth." *The Quarterly Journal of Economics*: 681-712.

[26] Moll-Murata Christine, 2008, "Chinese Guilds from the Seventeenth to the Twentieth Centuries: An Overview". *International Review of Social History*, Vol 53 (supplement S16) Dec, pp 213-247

[27] Munshi, Kaivan. 2003: "Networks in the modern economy: Mexican migrants in the US labor market," *Quarterly Journal of Economics*, 118, 549–599.

[28] Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. "The Allocation of Talent: Implications for Growth." *The Quarterly Journal of Economics* 106.2 (1991): 503-30.

[29] Ni, Xiaole, 2015, "China's Research and Development Spend", *Nature*, 30 April, VOL 520.

[30] Olken, Benjamin A. "Corruption and the costs of redistribution: Micro evidence from Indonesia." *Journal of Public Economics* 90.4 (2006): 853-870.

[31] Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike three: discrimination, incentives, and evaluation. *The American Economic Review*, 1410-1435.

[32] Romer, Paul M. "Increasing returns and long-run growth." *The Journal of Political Economy* (1986): 1002-1037.

[33] Shi Yigong and Yi Rao, 2010, "China's Research Culture", *Science*, 3 September: Vol. 329 no. 5996, p. 1128. DOI: 10.1126/science.1196916

[34] Siegel Jordan, 2007, Contingent Political Capital and International Alliances: Evidence from South Korea, *Administrative Science Quarterly*, December vol. 52 no. 4 621-666

[35] Young, Alwyn. "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period." *Journal of Political Economy* 111.6 (2003): 1220-1261.

[36] Zhu, Xiaodong. 2012. "Understanding China's Growth: Past, Present and Future." *Journal of Economic Perspectives* 26(4), pp. 103–124.

[37] Zinovyeva, Natalia, and Manuel Bagues. 2015. "The Role of Connections in Academic Promotions." *American Economic Journal: Applied Economics*, 7(2): 264-92.

Table 1a: Summary Statistics, Full Sample

Varible Name	Mean	StdDev	Observations
$CommitteeTie$	0.099	0.299	4922
$NonCommitteeTie$	0.334	0.472	4922
$Elected$	0.143	0.350	4922
$FirstStage$	0.404	0.491	4358
$Elected \vert FirstStage = 1$	0.338	0.473	1760
$log(1 + HIndex)$	1.597	1.256	4922
$HasPhD$	0.457	0.498	4922
Age	58.384	8.849	4845
$PoliticallyConnected$	0.048	0.214	4922
$Dean$	0.403	0.491	4922
$Committee_CollegeTie$	0.246	0.431	4922
$NonCommittee_CollegeTie$	0.448	0.497	4922

Notes: $CommitteeTie$ is an indicator variable denoting that the candidate shared his hometown with a standing committee member in the year of nomination. $NonCommitteeTie$ denotes a hometown connection to a department fellow not on the standing committee. $Elected$ denotes that a candidate was elected to the CAS or CAE in year y. $FirstStage$ denotes that a candidate was successful in passing through the first stage of selection to the CAS or CAE in year y. $HIndex$ is a the measure of a candidate's publishing record in the year of application, as developed by Hirsch (2005). $HasPhD$ is an indicator variable denoting that the candidate holds a Ph.D. Age is the candidate's age in the year of nomination. $PoliticallyConnected$ denotes candidates with government rank of vice-minister or higher, and $Dean$ denotes a candidate holding an academic position of dean or higher. $Committee_CollegeTie$ denotes that the candidate shared his undergraduate institution with a standing committee member in the year of nomination. $NonCommittee_CollegeTie$ denotes that the candidate shared his undergraduate institution with a department fellow not on the standing committee. See text for further details on variable construction. Table 1A provides summary statistics on the overall sample. Tables 1B and 1C compare the mean characteristics of applicants by committee ties and non-committee ties respectively.

Table 1b: Summary Statistics, by Committee Ties

	$CommitteeTie = 1$		$CommitteeTie = 0$		Difference	
Varible Name	Mean	StDev	Mean	StDev	Difference	t-statistic
$Elected$	0.197	0.398	0.137	0.344	0.060	3.627
$FirstStage$	0.388	0.488	0.406	0.491	-0.018	-0.694
$Elected \vert FirstStage = 1$	0.497	0.502	0.322	0.467	0.175	4.493
$log(1 + HIndex)$	1.601	1.204	1.596	1.262	0.005	0.076
$HasPhD$	0.398	0.490	0.463	0.499	-0.065	-2.726
Age	59.698	8.458	58.239	8.880	1.459	3.432
$PoliticallyConnected$	0.035	0.184	0.049	0.217	-0.014	-1.419
$Dean$	0.372	0.484	0.406	0.491	-0.035	-1.480
$Committee_CollegeTie$	0.402	0.491	0.229	0.420	0.174	8.504
$NonCommittee_CollegeTie$	0.544	0.499	0.437	0.496	0.107	4.523

Table 1c: Summary Statistics, by Non-Committee Ties

	$NonCommitteeTie = 1$		$NonCommitteeTie = 0$		Difference	
Varible Name	Mean	StDev	Mean	StDev	Difference	t-statistic
$Elected$	0.155	0.362	0.136	0.343	0.019	1.799
$FirstStage$	0.396	0.489	0.408	0.491	-0.011	-0.722
$Elected \vert FirstStage = 1$	0.379	0.485	0.318	0.466	0.061	2.530
$log(1 + HIndex)$	1.540	1.205	1.625	1.280	-0.085	-2.239
$HasPhD$	0.389	0.488	0.491	0.500	-0.101	-6.763
Age	59.939	8.597	57.608	8.873	2.330	8.702
$PoliticallyConnected$	0.049	0.217	0.047	0.212	0.002	0.321
$Dean$	0.386	0.487	0.412	0.492	-0.026	-1.759
$Committee_CollegeTie$	0.324	0.468	0.207	0.405	0.117	9.057
$NonCommittee_CollegeTie$	0.535	0.499	0.404	0.491	0.132	8.826

Notes: $CommitteeTie$ is an indicator variable denoting that the candidate shared his hometown with a standing committee member in the year of nomination. $NonCommitteeTie$ denotes a hometown connection to a department fellow not on the standing committee. $Elected$ denotes that a candidate was elected to the CAS or CAE in year y. $FirstStage$ denotes that a candidate was successful in passing through the first stage of selection to the CAS or CAE in year y. $HIndex$ is a the measure of a candidate's publishing record in the year of application, as developed by Hirsch (2005). $HasPhD$ is an indicator variable denoting that the candidate holds a Ph.D. Age is the candidate's age in the year of nomination. $PoliticallyConnected$ denotes candidates with government rank of vice-minister or higher, and $Dean$ denotes a candidate holding an academic position of dean or higher. $Committee_CollegeTie$ denotes that the candidate shared his undergraduate institution with a standing committee member in the year of nomination. $NonCommittee_CollegeTie$ denotes that the candidate shared his undergraduate institution with a department fellow not on the standing committee. See text for further details on variable construction. Table 1A provides summary statistics on the overall sample. Tables 1B and 1C compare the mean characteristics of applicants by committee ties and non-committee ties respectively.

Table 2: Standing Committee Hometown Ties and Candidate Election Rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.049**	0.048**	0.048**	0.047*	0.064***	0.052	0.045*
	(0.020)	(0.020)	(0.020)	(0.024)	(0.025)	(0.032)	(0.025)
NonCommitteeTie	0.005	0.007	0.015	-0.018	-0.002	-0.010	0.017
	(0.011)	(0.011)	(0.011)	(0.017)	(0.014)	(0.020)	(0.013)
$log(1 + HIndex)$	0.031***	0.028***		0.025***	0.033***	0.034***	0.025***
	(0.005)	(0.005)		(0.007)	(0.007)	(0.009)	(0.006)
HasPhD	0.025**	0.023*		0.019	0.020	0.032	0.020
	(0.012)	(0.012)		(0.016)	(0.017)	(0.024)	(0.014)
Dean	0.006	0.010		0.009	0.002	0.021	0.005
	(0.010)	(0.010)		(0.013)	(0.014)	(0.021)	(0.012)
PoliticallyConnected	0.026	0.033		0.061**	0.041	0.130	0.028
	(0.023)	(0.024)		(0.029)	(0.032)	(0.136)	(0.024)
$log(Age)$	0.110***	0.119***		0.192***	0.231***	0.136**	0.118***
	(0.037)	(0.037)		(0.048)	(0.051)	(0.067)	(0.045)
Committee_CollegeTie	0.018	0.016		0.025	-0.009	0.038	0.003
	(0.014)	(0.014)		(0.016)	(0.019)	(0.024)	(0.017)
NonCommittee_CollegeTie	0.014	0.016		0.026*	-0.010	0.006	0.020
	(0.011)	(0.011)		(0.014)	(0.018)	(0.020)	(0.014)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4845	4845	4922	4844	4656	1800	3045
R-Squared	.0246	.0339	.0238	.176	.151	.0334	.023

Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS and CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1 + Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Standing Committee Hometown Ties and Candidate Success, in the First Stage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*FirstStage*			
CommitteeTie	-0.015	-0.019	-0.020	-0.010	-0.009	0.017	-0.037
	(0.028)	(0.028)	(0.028)	(0.033)	(0.035)	(0.050)	(0.034)
NonCommitteeTie	-0.015	-0.013	0.000	-0.053**	-0.020	0.007	-0.020
	(0.019)	(0.019)	(0.019)	(0.027)	(0.023)	(0.039)	(0.022)
$log(1 + HIndex)$	0.068***	0.061***		0.056***	0.068***	0.062***	0.060***
	(0.008)	(0.009)		(0.011)	(0.011)	(0.017)	(0.011)
HasPhD	0.063***	0.057**		0.056**	0.056*	0.024	0.067***
	(0.022)	(0.023)		(0.026)	(0.029)	(0.047)	(0.026)
Dean	-0.014	-0.011		-0.019	-0.029	-0.047	0.003
	(0.018)	(0.018)		(0.021)	(0.022)	(0.036)	(0.021)
PoliticallyConnected	0.000	0.014		0.032	0.001	0.029	0.015
	(0.039)	(0.040)		(0.044)	(0.046)	(0.203)	(0.041)
$log(Age)$	0.147**	0.169**		0.214***	0.284***	-0.010	0.256***
	(0.068)	(0.067)		(0.080)	(0.088)	(0.131)	(0.079)
Committee_CollegeTie	0.010	0.009		0.010	-0.041	-0.003	0.011
	(0.022)	(0.022)		(0.025)	(0.029)	(0.040)	(0.027)
NonCommittee_CollegeTie	0.051***	0.052***		0.065***	0.004	0.011	0.072***
	(0.019)	(0.019)		(0.022)	(0.028)	(0.036)	(0.023)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4285	4285	4358	4284	4110	1240	3045
R-Squared	.0457	.0652	.0442	.27	.247	.0417	.0434

Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS during 2003-2011 and the CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i made it through the first stage of candidate selection to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1 + Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Standing Committee Hometown Ties and Candidate Election Rates, Conditional on First Stage Success

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.151***	0.153***	0.156***	0.134**	0.159***	0.119*	0.177***
	(0.040)	(0.043)	(0.043)	(0.058)	(0.056)	(0.065)	(0.056)
NonCommitteeTie	0.031	0.037	0.044*	-0.012	0.016	-0.032	0.075**
	(0.024)	(0.026)	(0.025)	(0.042)	(0.035)	(0.042)	(0.033)
$log(1 + HIndex)$	0.018*	0.016		0.016	0.022	0.023	0.014
	(0.010)	(0.011)		(0.016)	(0.016)	(0.019)	(0.015)
HasPhD	0.015	0.012		0.022	-0.003	0.045	-0.004
	(0.028)	(0.030)		(0.043)	(0.042)	(0.050)	(0.037)
Dean	0.032	0.035		0.061*	0.029	0.074*	0.015
	(0.023)	(0.025)		(0.035)	(0.034)	(0.039)	(0.032)
PoliticallyConnected	0.070	0.081		0.155**	0.136	0.240	0.067
	(0.053)	(0.056)		(0.070)	(0.083)	(0.237)	(0.057)
$log(Age)$	0.158*	0.179**		0.400***	0.346***	0.293**	0.122
	(0.085)	(0.091)		(0.139)	(0.126)	(0.138)	(0.122)
Committee_CollegeTie	0.028	0.021		0.050	0.007	0.074	0.001
	(0.029)	(0.031)		(0.042)	(0.045)	(0.053)	(0.039)
NonCommittee_CollegeTie	-0.021	-0.020		-0.016	-0.035	-0.037	-0.020
	(0.024)	(0.025)		(0.038)	(0.042)	(0.040)	(0.033)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	1742	1742	1760	1742	1702	658	1084
R-Squared	.0366	.0528	.0464	.31	.249	.0665	.0523

Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS during 2003-2011 and the CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1 + Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.

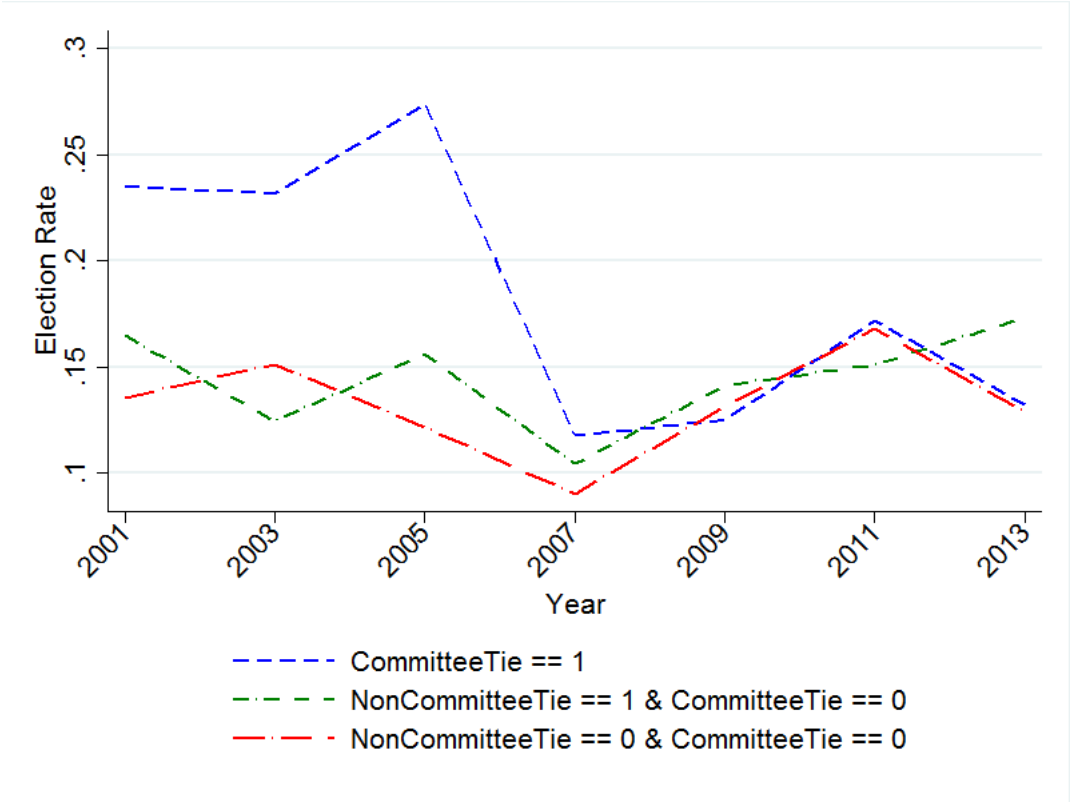
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Interaction of Candidate Quality and Hometown Ties in Predicting Candidate Success

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			*Elected*			
$log(1+HIndex)$	0.031***	0.025***	0.028***	0.032***	0.024***	0.026***
	(0.005)	(0.006)	(0.007)	(0.007)	(0.008)	(0.009)
$CommitteeTie$	0.101***	0.119***	0.120***	0.125***	0.138***	0.139***
	(0.033)	(0.034)	(0.034)	(0.038)	(0.039)	(0.039)
$NonCommitteeTie$	0.006	-0.031*	-0.031*	-0.023	-0.055**	-0.055**
	(0.011)	(0.017)	(0.017)	(0.017)	(0.023)	(0.024)
$CommitteeTie * log(1+HIndex)$	-0.032**	-0.043***	-0.043***	-0.049***	-0.055***	-0.056***
	(0.015)	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)
$NonCommitteeTie * log(1+HIndex)$		0.023**	0.023**		0.021*	0.020*
		(0.009)	(0.010)		(0.012)	(0.012)
Department X Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes	Yes	Yes
Controls Included?	Yes	Yes	Yes	Yes	Yes	Yes
College Interactions Incl?			Yes			Yes
Observations	4845	4845	4845	4844	4844	4844
R-Squared	.035	.0362	.0368	.178	.179	.179

Notes: Standard Errors clustered by candidate in all regressions. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. $CommitteeTie$ is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. $NonCommitteeTie$ denotes a hometown connection to a department fellow not on the standing committee. $log(1+Hindex)$ is self-explanatory. See the text for further details on variable construction. All specifications include the control variables used in Tables 2 - 4, and Columns 3 and 6 also include the interaction of $log(1+HIndex)$ with $Committee_CollegeTie$ and $NonCommittee_CollegeTie$. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Figure 1: Hometown Ties and Candidate Election Rates over Time



Notes: Each line provides the fraction of nominees elected to the CAS/CAE, disaggregated based on whether they have ties to fellows in the department of their nomination. The connections that characterize each group are provided in the figure legend.

Figure 2a: Candidate quality, hometown ties, and selection

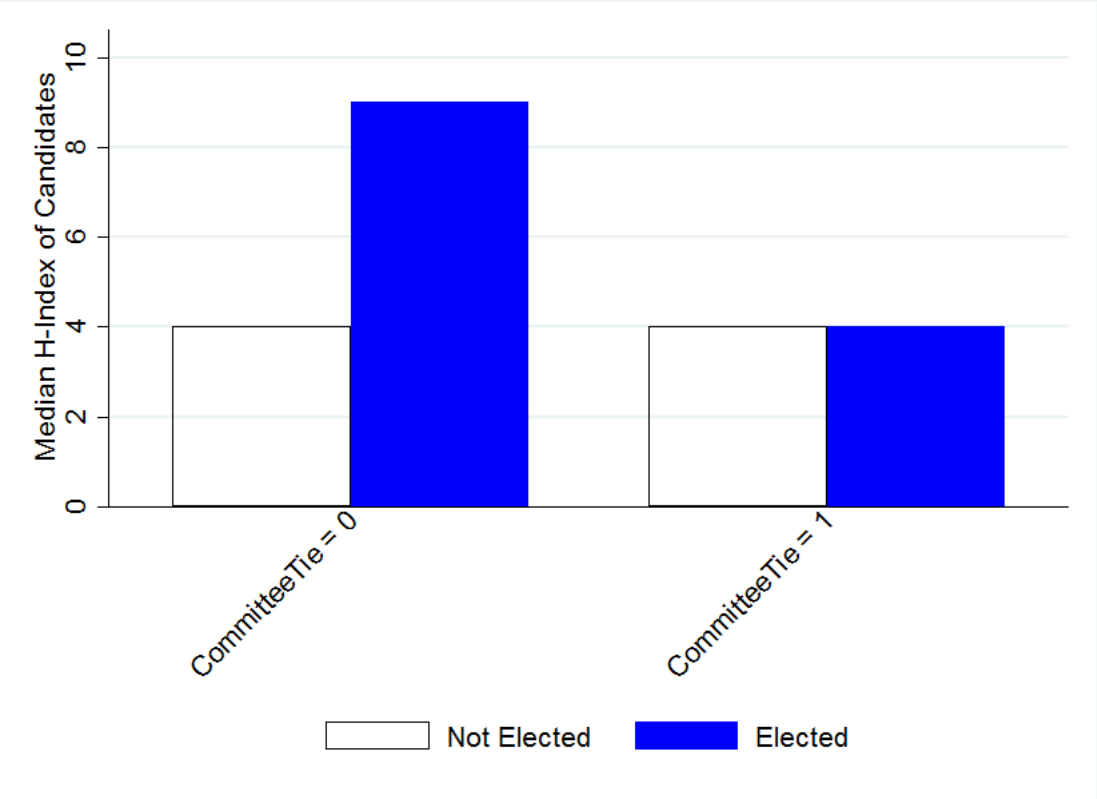


Figure 2b: Candidate quality, hometown ties, and first stage selection

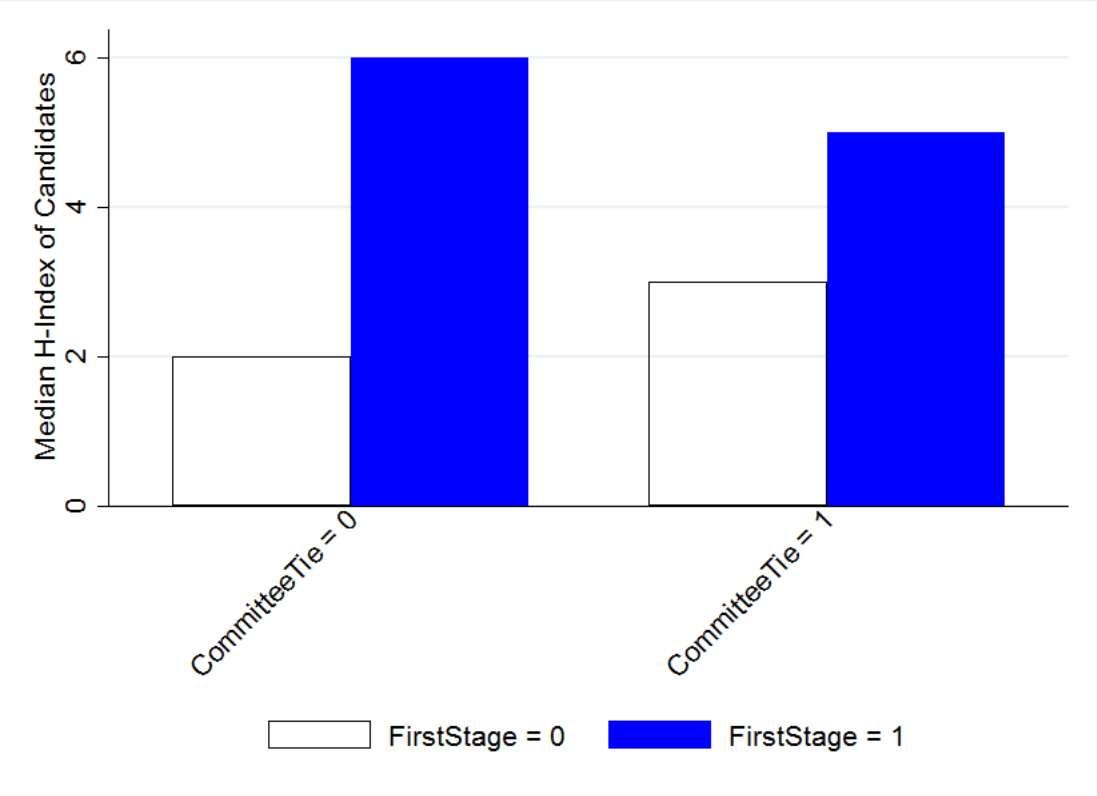
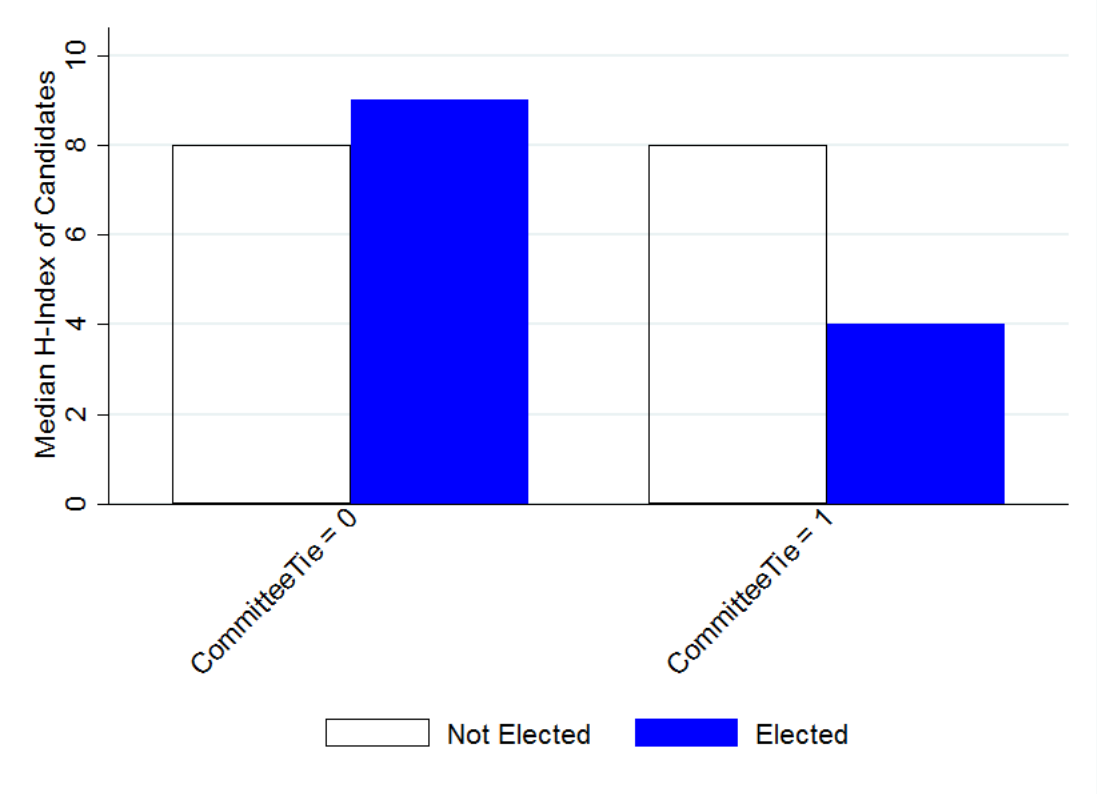


Figure 2c: Candidate quality, hometown ties, and second stage selection



Notes: Each panel of figure 2 provides a comparison of median H-Indexes of selected versus non-selected candidates, disaggregated by *CommitteeTie*. Panel A provides a comparison based on elected versus non-elected candidates (i.e., combining both stages of selection. Panel B compares candidates based on first-stage selection. Panel C compares elected versus non-elected candidates, limiting the sample to those candidates who succeeded in the first stage of selection.

Table A1: Standing Committee Hometown Ties and Candidate Election Rates, Placebo Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.049**	0.048**	0.047**	0.044*	0.064**	0.056*	0.043*
	(0.020)	(0.020)	(0.020)	(0.024)	(0.025)	(0.033)	(0.026)
NonCommitteeTie	0.005	0.007	0.013	-0.021	-0.001	-0.006	0.014
	(0.012)	(0.012)	(0.012)	(0.018)	(0.015)	(0.022)	(0.014)
CommitteeTie_Placebo	0.002	0.002	0.005	-0.039**	-0.002	-0.012	0.008
	(0.011)	(0.012)	(0.011)	(0.019)	(0.014)	(0.024)	(0.013)
$log(1 + HIndex)$	0.031***	0.028***		0.026***	0.033***	0.034***	0.025***
	(0.005)	(0.005)		(0.007)	(0.007)	(0.009)	(0.006)
HasPhD	0.025**	0.023*		0.020	0.020	0.032	0.020
	(0.012)	(0.012)		(0.016)	(0.017)	(0.024)	(0.014)
Dean	0.006	0.009		0.009	0.002	0.021	0.005
	(0.010)	(0.010)		(0.013)	(0.014)	(0.021)	(0.012)
PoliticallyConnected	0.026	0.034		0.059**	0.041	0.127	0.028
	(0.023)	(0.024)		(0.029)	(0.032)	(0.137)	(0.024)
$log(Age)$	0.109***	0.118***		0.196***	0.232***	0.139**	0.116**
	(0.037)	(0.037)		(0.048)	(0.051)	(0.068)	(0.045)
Committee_CollegeTie	0.018	0.016		0.025	-0.009	0.039	0.003
	(0.014)	(0.014)		(0.016)	(0.019)	(0.024)	(0.017)
NonCommittee_CollegeTie	0.014	0.016		0.025*	-0.010	0.006	0.020
	(0.011)	(0.011)		(0.014)	(0.018)	(0.020)	(0.014)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4845	4845	4922	4844	4656	1800	3045
R-Squared	.0246	.0339	.0239	.177	.151	.0336	.0231

Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS and CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1 + Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction. $CommitteeTie_Placebo$ denotes a hometown connection to a standing committee member *not* in the candidate's department.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A2: Standing Committee Hometown Ties and Candidate Election Rates, including age cohort fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.049**	0.047**	0.049**	0.046*	0.063**	0.050	0.045*
	(0.019)	(0.020)	(0.020)	(0.024)	(0.025)	(0.032)	(0.025)
NonCommitteeTie	0.006	0.009	0.012	-0.017	-0.000	-0.008	0.018
	(0.011)	(0.011)	(0.011)	(0.017)	(0.014)	(0.020)	(0.013)
$log(1 + HIndex)$	0.032***	0.028***		0.026***	0.034***	0.033***	0.025***
	(0.005)	(0.005)		(0.007)	(0.007)	(0.009)	(0.006)
HasPhD	0.021*	0.018		0.012	0.013	0.025	0.015
	(0.012)	(0.012)		(0.016)	(0.017)	(0.024)	(0.014)
Dean	0.005	0.008		0.007	-0.000	0.019	0.004
	(0.010)	(0.010)		(0.013)	(0.014)	(0.021)	(0.012)
PoliticallyConnected	0.026	0.034		0.062**	0.041	0.113	0.030
	(0.023)	(0.024)		(0.029)	(0.032)	(0.137)	(0.024)
Committee_CollegeTie	0.019	0.017		0.027*	-0.009	0.037	0.006
	(0.014)	(0.014)		(0.016)	(0.019)	(0.024)	(0.017)
NonCommittee_CollegeTie	0.015	0.017		0.027*	-0.011	0.011	0.020
	(0.011)	(0.011)		(0.014)	(0.018)	(0.020)	(0.014)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4845	4845	4845	4844	4656	1800	3045
R-Squared	.0268	.0362	.0277	.178	.153	.0396	.0251

Notes: Standard Errors clustered by candidate in all regressions, and all regressions include 5 year age cohort fixed effects. The sample in columns 1 - 5 includes all applicants to the CAS and CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1+Hindex)$ and $HasPhD$ are self-explanatory. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3: Employer Ties and Candidate Election Rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.048**	0.048**	0.048**	0.048**	0.058**	0.050	0.046*
	(0.020)	(0.020)	(0.020)	(0.024)	(0.025)	(0.032)	(0.026)
NonCommitteeTie	0.003	0.005	0.011	-0.020	-0.003	-0.013	0.015
	(0.011)	(0.011)	(0.011)	(0.017)	(0.015)	(0.020)	(0.013)
$log(1 + HIndex)$	0.026***	0.024***		0.022***	0.025***	0.030***	0.022***
	(0.005)	(0.005)		(0.007)	(0.008)	(0.009)	(0.007)
HasPhD	0.024**	0.023*		0.019	0.015	0.033	0.019
	(0.012)	(0.012)		(0.016)	(0.019)	(0.024)	(0.014)
Dean	0.008	0.010		0.010	0.024	0.025	0.005
	(0.010)	(0.010)		(0.013)	(0.015)	(0.021)	(0.012)
PoliticallyConnected	0.031	0.036		0.064**	0.050	0.135	0.030
	(0.023)	(0.024)		(0.029)	(0.039)	(0.137)	(0.024)
$log(Age)$	0.117***	0.124***		0.194***	0.217***	0.156**	0.114**
	(0.037)	(0.037)		(0.048)	(0.054)	(0.067)	(0.045)
Committee_CollegeTie	0.012	0.011		0.021	0.009	0.029	0.002
	(0.014)	(0.014)		(0.016)	(0.018)	(0.024)	(0.017)
NonCommittee_CollegeTie	0.011	0.013		0.022	0.021	0.005	0.017
	(0.011)	(0.011)		(0.014)	(0.016)	(0.020)	(0.014)
Committee_EmployerTie	0.034**	0.030**	0.033**	0.017	0.024	0.068***	-0.003
	(0.015)	(0.015)	(0.015)	(0.018)	(0.022)	(0.023)	(0.020)
NonCommittee_EmployerTie	0.023*	0.020	0.032***	0.029*	-0.027	0.003	0.031**
	(0.012)	(0.013)	(0.012)	(0.015)	(0.019)	(0.022)	(0.015)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
Employer FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4845	4845	4922	4844	4845	1800	3045
R-Squared	.0277	.0363	.0284	.178	.19	.0395	.0246

Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS and CAE during 2001-2013; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member in his department of nomination. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. *Committee_EmployerTie* denotes a candidate who shares an employer with a standing committee member in his department of nomination. *NonCommittee_CollegeTie* denotes a candidate who shares an employer with a department fellow not on the standing committee. $log(1 + Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A4: Connections and Selection, pre vs post-2007

	(1)	(2)	(3)	(4)
$CommitteeTie$	0.088***	-0.008	0.231***	0.132***
	(0.028)	(0.037)	(0.060)	(0.042)
$NonCommitteeTie$	0.012	0.013	0.031	-0.009
	(0.016)	(0.026)	(0.037)	(0.022)
$log(1 + HIndex)$	0.036***	0.054***	0.041**	0.033***
	(0.008)	(0.013)	(0.018)	(0.010)
$HasPhD$	0.005	0.029	-0.015	0.004
	(0.011)	(0.020)	(0.026)	(0.011)
$Dean$	0.006	-0.016	0.028	0.007
	(0.010)	(0.018)	(0.024)	(0.010)
$PoliticallyConnected$	0.033	0.010	0.080	0.034
	(0.023)	(0.040)	(0.054)	(0.024)
$Committee_CollegeTie$	0.017	0.010	0.024	0.017
	(0.014)	(0.022)	(0.031)	(0.014)
$NonCommittee_CollegeTie$	0.017	0.057***	-0.016	0.017
	(0.011)	(0.019)	(0.025)	(0.011)
$CommitteeTie * I(Year \geq 2007)$	-0.091**	-0.034	-0.164*	-0.062
	(0.039)	(0.054)	(0.090)	(0.063)
$NonCommitteeTie * I(Year \geq 2007)$	-0.000	-0.038	0.021	-0.044
	(0.022)	(0.036)	(0.051)	(0.033)
$log(1 + HIndex) * I(Year \geq 2007)$	-0.010	0.018	-0.038	-0.014
	(0.011)	(0.017)	(0.023)	(0.012)
$CommitteeTie * log(1 + HIndex)$				-0.034
				(0.024)
$NonCommitteeTie * log(1 + HIndex)$				0.016
				(0.015)
$CommitteeTie * log(1 + HIndex) * I(Year \geq 2007)$				-0.002
				(0.032)
$NonCommitteeTie * log(1 + HIndex) * I(Year \geq 2007)$				0.017
				(0.020)
Full Sample	Yes	Yes		Yes
FirstStage = 1 Sample			Yes	
Dep Var Elected	Yes		Yes	Yes
Dep Var FirstStage		Yes		
Observations	4922	4358	1760	4922
R-Squared	.0341	.0656	.0536	.0362

Notes: Standard Errors clustered by candidate in all regressions. The dependent variable in columns 1, 3, and 4 is *Elected* while in column 2 it is $FirstStage$. We use the full sample in columns 1, 2, and 4, while in column 3 the sample includes only $FirstStage = 1$ candidates. $CommitteeTie$ is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. $NonCommitteeTie$ denotes a connection to a department fellow not on the standing committee. $Dean$ denotes a candidate holding an academic position of dean or higher, and $PoliticallyConnected$ denotes candidates with government rank of vice-minister or higher. $log(1 + Hindex)$ and $log(Age)$ are self-explanatory. $I(Year \geq 2007)$ denotes observations from application years 2007 and later. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A5: Standing Committee Hometown Ties and Candidate Election Rates, Subsample with *FirstStage* Data (i.e., excl CAS 2001,2013)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable				*Elected*			
CommitteeTie	0.053**	0.051**	0.052**	0.048*	0.073***	0.061	0.045*
	(0.021)	(0.022)	(0.021)	(0.026)	(0.028)	(0.040)	(0.025)
NonCommitteeTie	0.007	0.009	0.016	-0.023	-0.000	-0.014	0.017
	(0.012)	(0.012)	(0.012)	(0.018)	(0.015)	(0.025)	(0.013)
$log(1 + HIndex)$	0.031***	0.027***		0.025***	0.033***	0.033***	0.025***
	(0.005)	(0.006)		(0.007)	(0.007)	(0.011)	(0.006)
HasPhD	0.025**	0.023*		0.022	0.020	0.037	0.020
	(0.013)	(0.013)		(0.016)	(0.018)	(0.029)	(0.014)
Dean	0.006	0.009		0.011	0.003	0.022	0.005
	(0.011)	(0.011)		(0.013)	(0.014)	(0.023)	(0.012)
PoliticallyConnected	0.025	0.033		0.060**	0.041	0.158	0.028
	(0.023)	(0.024)		(0.029)	(0.033)	(0.154)	(0.024)
$log(Age)$	0.113***	0.125***		0.178***	0.229***	0.165**	0.118***
	(0.039)	(0.039)		(0.050)	(0.054)	(0.080)	(0.045)
Committee_CollegeTie	0.015	0.012		0.018	-0.014	0.039	0.003
	(0.015)	(0.015)		(0.017)	(0.021)	(0.030)	(0.017)
NonCommittee_CollegeTie	0.009	0.011		0.018	-0.014	-0.014	0.020
	(0.012)	(0.012)		(0.015)	(0.019)	(0.023)	(0.014)
Department; Year FEs	Yes						
Department X Year FEs		Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4285	4285	4358	4284	4110	1240	3045
R-Squared	.0227	.0333	.0235	.187	.162	.0388	.023

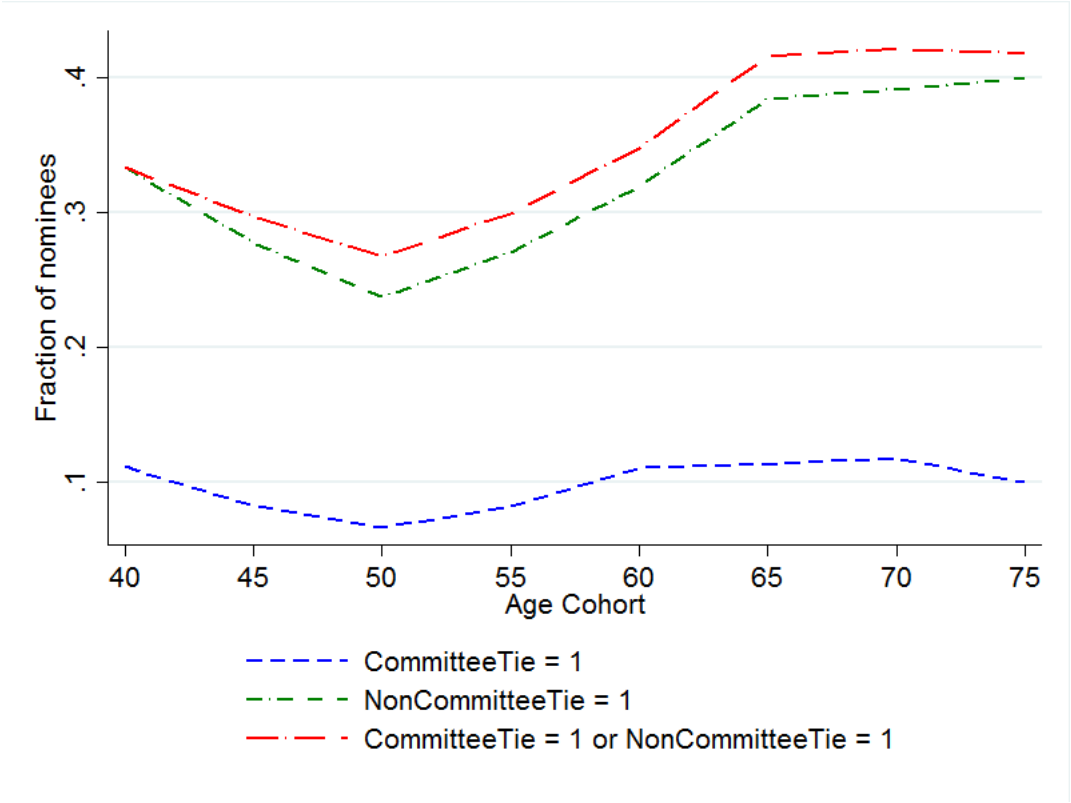
Notes: Standard Errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all applicants to the CAS and CAE during 2001-2013 for the CAE and 2003-2011 for the CAS; columns 6 and 7 provide results on the CAS and CAE separately. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member. *NonCommittee_CollegeTie* denotes a candidate with a college connection to a fellow not on the standing committee. $log(1+Hindex)$, *HasPhD*, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A6: Committee Hometown Ties and Future Productivity of Fellows

	(1)	(2)	(3)	(4)
Dependent Variable	$log(1+(HIndexin2014))$			
Elected	0.042**	0.066	0.037*	0.037*
	(0.019)	(0.081)	(0.020)	(0.020)
$log(1+HIndex)$	0.839***	0.796***	0.843***	0.839***
	(0.013)	(0.040)	(0.014)	(0.013)
CommitteeTie	-0.032	0.000	0.000	-0.038
	(0.028)	(.)	(.)	(0.030)
NonCommitteeTie	-0.009	-0.046	-0.009	-0.009
	(0.021)	(0.058)	(0.023)	(0.021)
HasPhD	0.085***	0.080	0.080***	0.085***
	(0.024)	(0.072)	(0.026)	(0.024)
Dean	0.059***	0.098	0.055***	0.059***
	(0.019)	(0.064)	(0.020)	(0.019)
PoliticallyConnected	0.038	0.036	0.039	0.038
	(0.048)	(0.119)	(0.051)	(0.048)
$log(Age)$	-0.383***	-0.341	-0.389***	-0.383***
	(0.076)	(0.260)	(0.080)	(0.076)
Committee_CollegeTie	-0.004	0.122*	-0.016	-0.004
	(0.023)	(0.065)	(0.025)	(0.023)
NonCommittee_CollegeTie	-0.019	-0.038	-0.017	-0.019
	(0.019)	(0.061)	(0.021)	(0.019)
$CommitteeTie * Elected$				0.034
				(0.069)
Full Sample	Yes			Yes
Same City Fellow(s) on Committee == 1		Yes		
Same City Fellow(s) on Committee == 0			Yes	
Observations	4845	480	4365	4845
R-Squared	.851	.873	.851	.851

Notes: Standard Errors clustered by candidate in all regressions. The sample in all columns includes all applicants to the CAS and CAE during 2001-2013. The dependent variable in all columns is the logarithm of one plus the candidate's H-Index in 2014. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of application. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with government rank of vice-minister or higher. *Committee_CollegeTie* denotes a candidate that attended the same undergraduate institution as a standing committee member

Figure A1: Nominees' Connections to Fellows by Age Cohort



Notes: Each line provides the fraction of nominees, by age cohort, with connections to CAS/CAE fellows in the department of their nomination. The type of connection (on standing committee; off standing committee; either) is provided in the figure legend We use 5 year cohorts starting at the age of 40, and ending at 79, since very few candidates' ages lie outside of this range (particularly at the upper end).