Social Ties and Favoritism in Chinese Science

Raymond Fisman
Boston University
Jing Shi
RMIT College of Business
Yongxiang Wang
University of Southern California
Rong Xu
Renmin University of China and China Financial Policy Center *

March 22, 2017

*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, and China Financial Policy Center, 59 Zhongguancun St, Haidian, Beijing, China; email: xurong@ruc.edu.cn. We are grateful for comments and suggestions from Jesse Shapiro and six anonymous referees.We would also like to thank Harry DeAngelo, Murray Frank, David Hirshleifer, Jerry Hoberg, Christian Julliard, Danielle Li, John Matsusaka, and Jun Pan, as well as seminar participants at a number of institutions for useful comments. Wang thank the China National Natural Science Foundation (Grant Number: 71432008) and Xu thank the China National Natural 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.

Abstract

We study favoritism via hometown ties, a common source of favor exchange in China, in fellow selection of the Chinese Academies of Sciences and Engineering. Hometown ties to fellow selection committee members increase candidates' election probability by 39 percent, coming entirely from the selection stage involving an in-person meeting. Elected hometown-connected candidates are half as likely to have a high-impact publication as elected fellows without connections. CAS/CAE membership increases the probability of university leadership appointments and is associated with a US$9.5 million increase in annual funding for fellows' institutions, indicating that hometown favoritism has potentially large effects on resource allocation.

Keywords: knowledge production; social networks; resource misallocation

Economists have long considered the consequences of rent-seeking and the resultant allocative inefficiency for economic growth. Earlier work has focused on the misallocation of capital (human or physical) in production (e.g., Murphy et al. (1991), Acemoglu (1995)). However, 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. Consequently, 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 of distortions in 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 expenditure has grown at an annual rate of more than 20 percent in recent years, and China is currently second only to the U.S. in R&D spending (Ni (2015)).

Press accounts have argued that some of this expenditure has 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

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

to research resources. It also translates into direct material rewards, 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 totaling 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 CAS membership. He came up one vote short in his second attempt in 2009, despite never having 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 process 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, Yigong Shi and Yi Rao, 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] Furthermore, Shi and Rao 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." Shi and Rao (2010) thus raise concerns about possible misallocation of effort by scientists (lobbying versus research) and also about misallocation of resources across scientists (from good researchers to effective lobbyists). Moreover, their narrative 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 of the CAS and CAE during the 2001-2013

[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.

period, using a form of connections that plays a central role in Chinese society: hometown ties. We focus on hometown ties, or *laoxiang guanxi* in Chinese, because of their importance in the culture of favor-exchange (*guanxi*) 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 along disciplinary lines, such as chemistry, mathematics, and so forth).

We show that, during 2001-2013, the probability of a nominee to the CAS or CAE being elected as a fellow was 39 percent higher if he was connected according to our measure. This result is highly robust. It survives the inclusion of department-year fixed effects, city fixed effects, undergraduate college fixed effects, and employer fixed effects. Further, we find no effect of 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 from the candidate's own department. Finally, we do not find a robust effect of connections to fellows via a candidate's undergraduate institution or current employer, which suggests that our results are unlikely to result from "soft" information on candidate quality, which would likely be captured by shared education or employment.

We 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. First-stage selection is correlated instead with candidates' publication

[4]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 (Lauren et al. (2008, 2010)), and among politicians in the U.S. Congress (Cohen and Malloy (2014)).

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 can more easily influence voting behavior. In this second stage, publication records are not predictive of success.

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 nominees.

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 among connected candidates. Based on analyses of candidates' H-Indexes and other measures of research success, we find no support for the former, i.e., nominee quality is unrelated to connections. However, we find strong support for the latter prediction. For example, among candidates who are ultimately elected to the CAS/CAE, those with hometown ties are about half as likely to have had at least one "homerun" (100+ published citations) relative to candidates without such connections. We show that this difference comes primarily from positive selection among unconnected nominees in the election process. Our calculations indicate that a prohibition on fellows evaluating candidates from their hometown would increase the fraction of elected fellows with a homerun by 2.7 percentage points; moving from a hypothetical scenario where all candidates have hometown ties to one where none do would increase the fraction of elected fellows with a homerun by nearly 20 percentage points.

In our final set of analyses, we show that election to the CAS/CAE more than doubles a scientist's probability of being appointed dean or president of a university, and that employ-

ment of CAS/CAE fellows is associated with an estimated US$ 9.5 million in incremental government funding for a fellow's institution. These findings on the greater power and resources that come with CAS/CAE election indicates that the favoritism that we document may have significant implications for the allocation of research resources.

There are two main limitations to the interpretation of our results. First, we only observe a single channel of favoritism, which makes it difficult to generate a decisive counterfactual estimate of what the quality of CAS/CAE membership would be in the absence of favoritism and rent-seeking (though in Section 3.2 we provide some discussion of this issue). Second, we cannot directly measure the impact of favoritism on ultimate scientific outcomes. We show that favoritism leads to lower-quality scientists, but it goes beyond the scope of our paper to quantify the full effect of favoritism on the allocation of scientific resources.

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., and finds that applicants connected to a reviewer via citation history are more likely to receive funding. Zinovyeva and Bagues (2015) find that professional ties to evaluators predicts academic promotions in Spain. Durante et al. (2014) find evidence that family ties play a role in academic hiring in Italy.[5]

Empirical 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 (see also Khandelwal et al. (2013) and Brandt et al. (2012)). Our study provides more direct evidence on the allocation mechanism that produces distortions (fa-

[5]Parsons et al. (2011) discuss how favoritism along ethnic lines can distort decisions in a very different setting, showing that Major League Baseball umpires make favorable calls toward players of their own ethnicity.

voritism via hometown ties), and focuses on the distinct domain of scientific enterprise.

Finally, we contribute to the literature on the distortionary effects of in-group favoritism. Most directly related to our paper, there is growing concern within China about the abuses of *guanxi* where, as we discuss in detail below, hometown ties play an important role (see Li (2011) for a general treatment). The distortions from in-group favoritism are a global concern, however: Burgess et al. (2015), for example, show that in Kenya districts that share the president's ethnicity receive twice the road building funds and quadruple the length of paved roads as do unconnected districts.

In the next section we provide background information on the Chinese Academies and describe the process for electing new fellows, as well as a discussion of the role of hometown ties in Chinese society. In Section 2 we discuss the datasets that we employ in our analyses. Section 3 provides our empirical analysis on the role of hometown ties in CAS/CAE selection, as well as the consequences of hometown favoritism. Section 4 concludes.

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 1000 CAS-affiliated sites throughout China. In addition to promoting science through its affiliated institutions, the CAS serves as an academic society, with CAS membership seen as the country's highest scientific 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 members in 2014): Engineering Management; Chemical, Metallurgical and Material Engineering; Mechanical and Vehicle Engineering; Energy and Mining Engineering; Civil and Hydraulic Engineering; Light Industry and Environmental Engineering; Information and Electronic Engineering; Medicine and Health Engineering; Agriculture; and Light Industry and Environmental Engineering. The last two were created from a split of a single department in 2006.

Beyond the honor of membership, fellows enjoy a number 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 fellows 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 little over US$30,000), a starting research budget of at least one million RMB, and a car and driver. By comparison, a standard full professor's salary is less than half that amount.[6] 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.[7]

This eagerness to attract fellows is in part due to the funding and connections that come with CAS/CAE membership. The academies - and by extension their members - direct the allocation of significant research resources. 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"

[6] A regular full professor receives no such perks, and could expect to receive a salary of less than US$15,000. See Altbach (2012) for estimates of faculty salaries in China.

[7] 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.

grants from the Ministry of Science and Technology, or MOST (which has a budget that in 2014 approached US$10 billion) often require CAS or CAE fellow recommendations.[8] Furthermore, 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.

Beyond these narrative examples of the resources controlled by the CAS and CAE as organizations, in Section 3.3 we provide more direct empirical evidence on the influence and power of fellows at the individual-level.

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 balance of membership based on subfields, sectors, and also regions.[9] Prior to 2008, standing committee members in CAS departments 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. For each academy, the committee is comprised of the directors of each department, the dean and vice-deans of the academy, and a few other fellows that are elected

[8]The following link provided 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.

[9]The CAS website provides details of the standing committee election process: http://history.casad.cas.cn/document.action?docid=11998 (downloaded October 11, 2016)

by a general vote at the biennial academy meeting. Elected committee members have four-year terms, which are nonrenewable in the CAS (renewable only once in the CAE). This committee is in charge of the daily administration 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.[10] 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 forward amongst those under its administration. For example, Peking University is administered by the Ministry of Education. So Peking University may put forward nominations to the Ministry of Education which will assess these candidates and those from other universities, then decide which university-affiliated candidates will receive formal nominations.[11]

Within each department, selection among 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 written evaluations of applications along with a yes-no vote. Every member of the

[10] See the CAS bylaws for details on the current selection process http://english.casad.cas.cn/Ab/Re/ (downloaded October 11, 2016)

[11] 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 Agriculture, 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).

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 eliminate about 40 percent of the initial pool of candidates. We refer to this winnowing as the first selection stage.

The second stage begins with an evaluation of the remaining candidates 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 half of voting fellows to be elected).

This process gives standing committee members considerable sway in the selection of fellows. In the first stage, they assign candidates' applications to fellows within their departments for initial review. In 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 to exercise social pressure in the final in-person vote. Finally, while non-standing committee members can skip the biennial meeting at which selection takes place, standing committee members are required to attend: as department leaders their presence at the meetings is mandatory.[12] The election rules at the CAE are very similar to those of the CAS, except for minor differences. Most notably,

[12]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.

candidates who make it past the first stage are required to give 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 anthropologist Leo Douw in his introduction to a book-length treatment of the topic,"[t]he cultivation of hometown ties is part and parcel 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 Xiao-Ping and Chen (2004) observe, hometown ties are among the most common and distinctive bases for *guanxi* to build upon.

There is a literature too vast to survey here that examines the origins of *laoxiang guanxi*, and also documents its many roles 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 among migrant communities within China and also among the global Chinese diaspora.[13]

These connections have led to favor exchange that has been explicitly censured by Chinese government officials. In early 2015, a director at the Central Commission for Discipline Inspection, China's highest anti-corruption authority, expressed his concerns about the culture

[13]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, Freedman (1960); Crissman (1967), and Kerri (1976).

of favor-exchange that had emerged around hometown networks, worrying that government officials maintained their hometown ties solely for the purpose of building profitable connections to businesses or securing promotion.[14] 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.

Hometown ties have been implicated in corruption of the CAS/CAE selection process that is our focus. In October 2013, Sciencenet, a publication cosponsored 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.[15] 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 it came to light that some of his research contributions were fabricated or written by others, leading to his arrest for corruption in 2012.

2 Data

Candidate information for both the CAS and the 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 and second stages of 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

[14] See http://www.hebgcdy.com/2015/0123/104692.html accessed on October 11, 2016.
[15] See http://news.sciencenet.cn/htmlnews/2013/10/283957.shtm accessed October 11, 2016.

the list 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 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."[16] The CAE sent us the data from 2001.

The nominee lists that 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. Nominees who fail in their first bid for membership of the CAS and CAE may be nominated again in subsequent years, so a single candidate i may appear in multiple years. We match candidates over time based on name, birthplace, and birth year. Of our final sample of candidates, 1663 (49.7 percent) are nominated only once, 915 (27.4 percent) are nominated twice, and 768 receive nominations three or more times. (11 candidates were nominated to departments in the CAS and CAE in a single year, but otherwise all candidate-year observations are distinct.)

Conditional on receiving a nomination after an initial failure, a candidate's success rate is much higher. For example, the success rate is 7.1 percent for all candidates in their first attempt, versus 13.6 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 most nominees through a combination of employer websites and 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 jurisdiction), and educational

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

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 their 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 the city of birth for 766 candidates out of a total of 3349. Of these, 259 candidates (20.7 percent of the total of 1251) were CAS candidates, while 507 (21.7 percent of the total of 2332) were CAE candidates. These candidates are necessarily excluded from our analysis.[17] 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. (In 79 percent of cases, a connected candidate has only a single hometown tie to the standing committee, in 17 percent of cases there are two ties, and in 4 percent of cases a connected candidate has 3 or more ties.) 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 generate variables that indicate that

[17]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 0.70).

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.[18] We show in Appendix Table A1 that our results are robust to including controls for these educational and professional ties.

Throughout our analysis, we wish to control for academic output. Our main measure is a candidate's H-index at the end of the relevant election year, obtained from Web of Science (Core Collection). 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. In this, we follow the innovations literature, which typically allows for several years' lag in measuring citations; see, for example, Aghion et al. (2013). One potential concern is that our forward-looking measure of publication impact at the time of election incorporates any positive treatment effect that CAS/CAE membership has on citations. In Appendix Tables A2 and A3, we show that our results are robust to using publications as a measure of research quality, which is not subject to this concern.

In many cases, there were multiple search results due to common names and/or candidates' use of initials rather than full names. In these cases, we also matched based on the author's affiliated institution and field of research.[19]

To account for the long right tail in the H-Index distribution, as well as the fact that 36 percent of CAE nominees (as well as 6 percent of CAS nominees) have an H-Index of zero,

[18]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.

[19]While shared names are common enough in Chinese, it would be rare to have such overlap for two individuals within the same institution and the same field of research. In practice, after filtering by name and affiliated institution we do not find any cases of a name/institution combination where there are publications across unrelated fields.

we use $log(1 + HIndex_{yi})$ as our main measure of the research productivity of candidate *i*. (See Online Appendix Figure OA1 for a histogram of H-Indexes for the sample overall, and for the CAS and CAE separately.) While there is no sufficient statistic for observable candidate quality, the H-Index is an accepted measure that captures both quantity of output and citation impact (Hirsch (2005)).

To assess the robustness of our results to alternative measures, we also collected data on candidates' total publications and total citations, and on "homerun" publications – those with over 100 published citations on Web of Science. We employ an indicator variable, $HasHomerun$, that captures whether a candidate had a homerun (100+ citation) paper at the time of nomination.[20]

As a final measure of academic credentials, we also include $Doctorate_{yi}$, an indicator variable denoting that a candidate holds a Ph.D. or equivalent degree such as an M.D. (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 (in practice dean or president) at his academic institution, while $PoliticalTie_{yi}$ captures whether the nominee is (or was) a vice-*Tingju* level (or above) government official, where a vice-*Tingju* level official holds the same rank as a city vice-mayor. In our context, politically influential candidates are typically former government officials (including some of very high rank; for example, the former Minister of Railways, Fu Zhihuan, was a candidate in 2001 and was elected as a fellow in that year) or high-ranking members of the military.

[20]We also collected data via the *China Academic Literature Network Publishing General Database* at CNKI on candidates' research records in Chinese scientific journals, including citations, publications, and H-Indexes. We found that none of these productivity measures led to greater success in election to the CAS/CAE: the Chinese H-Index is *negatively* correlated with election, significantly so for the CAS. This confirmed our prior belief that, for the most part, Chinese journals are not well-regarded by the scientific establishment. See Appendix Tables A2 and A3 for these results.

Table 1A provides summary statistics for the full sample, while Tables 1B and 1C disaggregate the data 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.

In Table 1A we see that hometown ties are relatively rare: $CommitteeTie = 1$ for 10.0 percent of candidates.[21] Additionally, we observe that the average candidate is 58.4 years old - a reminder that election to the academies is a late-career reward for past accomplishments.

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 5.9 percentage points in the fraction of nominees that are elected fellows between $CommitteeTie = 1$ candidates and $CommitteeTie = 0$ candidates (19.6 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 $CommitteeTie = 1$ candidates), 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 17.2 percentage points higher for $CommitteeTie = 1$ candidates (49.4, versus 32.2 for $CommitteTie = 0$ candidates). The difference in success rates based on non-committee ties is much more modest and, as we will see in our regression results in the following section, does 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 large numbers 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

[21] By department, the rate of hometown ties ranges from 3.6 percent in Engineering Management to 16.1 in Mathematics and Physics. The rate is 11.2 percent for the CAS overall versus 9.4 for the CAE.

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 connected 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 doctoral degrees is a direct result of this age gap, as 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 candidates under 50), while the rate is below 20 percent for candidates aged 60 and over. Once we control for age, the difference in the fraction of candidates with doctoral degrees in Tables 1B and 1C disappears (see Online Appendix Table OA1).

This still leaves the question of why there is an age difference based on connections in the first place. We can offer one speculative answer. 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 prevalent 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 who 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

3.1 Hometown ties and election to the CAS/CAE

We begin in Figure 1 by showing how the gap in election rates by hometown connection status varies over 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. This is 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.

There 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). We speculate that an important factor may have been the increase in 2007 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 nominees, as it plausibly made it more difficult for influential fellows to secure enough votes to gain approval for their favored candidates. There were other concurrent changes that might also have affected candidate

selection. Since 2007, 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 newsletter, the *CAS Bulletin*, and on the CAE's own website. It is likely that both the increased publicity and the changes in electoral rules were responses to public criticism of the CAS/CAE fellow 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-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. We compute standard errors that allow for clustering by candidate, since a single individual may apply multiple times.[22]

In column 1, we show the results of specification (1) including only $CommitteeTie$ and $NonCommitteeTie$ as covariates, along with department-year fixed effects. The coefficient on $CommitteeTie$ is 0.050, significant at the 1 percent level. In column 2 we add a number of covariates: $log(1 + Hindex)$ and $Doctorate$ to proxy for candidate quality; controls for academic and political 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 $log(Age)$. The coefficient on $CommitteeTie$ increases slightly to 0.053 (significant at the 1 percent level). Recall that the mean success rate of $CommitteeTie = 0$ candidates is 0.137, so these estimates imply that a hometown tie increases the probability of becoming a fellow by about 39 percent. In column 3 we

[22]We may also cluster at the level of the election (i.e., department-year), given the non-independence of votes received within a department in a given year. This generates slightly larger standard errors, though all our full sample results remain significant at least at the 10 percent level. See Online Appendix Table OA2.

use $HasHomerun$, an indicator variable denoting whether a nominee has at least one 100+ citation paper, as a readily interpretable measure of research quality. As with our H-Index measure on column 2, $HasHomerun$ is very significant (p-value < 0.001). The magnitude of its coefficient, 0.060, is about 15 percent greater than that of $CommitteeTie$, indicating that a hometown connection has an impact on selection that is roughly comparable to that of having a high impact journal publication.

In column 4 we provide our most rigorous specification which includes hometown fixed effects for each of the 424 municipalities (including county cities) with at least one candidate during our sample, as well as department-year fixed effects. (62 of these municipalities have within-city variation in $CommitteeTie$, though they tend to be larger municipalities and contain 54 percent of candidate observations in our sample.) 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 5 percent level (p-value of 0.047). In column 5 we include fixed effects for candidates' undergraduate institutions. This leads to a modest increase in the coefficient on $CommitteeTie$, to 0.070, with a standard error that is slightly higher than our main specification in column 2. 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.

Overall, the results in Table 2 indicate that the effect of hometown ties to committee members is quite robust, and distinct from other measures of connectedness. Across columns 1 - 5, we may reject equality of coefficients for $CommitteeTie$ and $NonCommitteeTie$ at least at the 10 percent level (at least at the 5 percent level if we allow for any individual-level controls). Furthermore, we do not find a consistent impact of connectedness via undergraduate institution, a tie that would more plausibly be a conduit for soft information on scientific ability.

In Appendix Table A1, we provide a series of further robustness checks that highlight both the robustness and distinctiveness of the impact of hometown ties on CAS/CAE election outcomes. In column 1 we present a 'falsification test' by including $CommitteeTie\_Placebo$, which captures hometown ties to standing committee members *not* in the candidate's department. These non-department standing committee ties have no effect on a candidate's election prospects, as one would expect given that election is conducted at the department level. In column 2 we show that our findings are robust to the inclusion of a full set of age cohort fixed effects. The point estimate on β_1 is near-identical to those in Table 2, with a comparable standard error. In column 3 we include $Committee\_EmployerTie$ and $NonCommittee\_EmployerTie$, which capture nominees' ties to fellows through their employer at the time of nomination, as well as employer fixed effects. As with undergraduate ties, we argue that these professional connections would more plausibly serve as a channel for soft information than hometown ties. In this specification, the coefficient on $CommitteeTie$ increases to 0.061 (and the coefficient on $NonCommitteeTie$ is slightly negative though insignificant). If soft information were the primary reason for $CommitteeTie$'s effect on selection, it is very surprising that neither school nor employer ties have any positive effect. In column 4 we verify that the differences across time observed in Figure 1 are statistically significant. When we add the interaction of $CommitteeTie$ and an indicator variable denoting election years later than 2007, we find that the direct effect of $CommitteeTie$ increases to 0.094, while the interaction is of near-equal magnitude and opposite sign. Finally, to emphasize the robustness of our results to alternative measures of research impact, in Appendix Table A2 we present specifications comparable to those in Table 2, column 2, with $log(1 + Publications)$, $log(1 + Citations)$,$log(1 + ChineseHIndex)$, $Publications$, and H-Index deciles as controls. The coefficient on $CommitteeTie$ is stable across all specifications and, apart from $log(1 + ChineseHIndex)$, all measures of research output are significant predictors of election.[23] Finally, it is possible to show how election probabilities shift when

[23]Because an individual may be nominated more than once, we may also run our analysis with candidate fixed effects. We present these results in Online Appendix Table OA3. The coefficients on *CommitteeTie*

a city has a fellow elected to the standing committee, or when a city's fellow steps down. There are 60 such transitions where we have at least one candidate that is nominated both pre and post-transition. Using these data, we obtain selection probabilities that are very much in line with our regression results: in a year when a city gets a new standing committee member, its nominees' election probability increases from 14.7 to 21.3 percent. When a standing committee tie is lost, the selection probability decreases from 18.3 to 12.8 percent. Owing to the shortness of our panel, we cannot provide 'event studies' for these transitions with more years before and after the transition.

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.[24] We present the results in Tables 3, where we include individual-level controls and department-year fixed effects in all specifications (we suppress the coefficients on control variables to conserve space; these coefficients, along with more extensive first and second stage results, may be found in Online Appendix Tables OA4 and OA5). In columns 1 and 2 we present the results for the first selection stage. There are two interesting patterns that emerge. First, the link between observable candidate quality and progressing past the first stage is quite strong. The coefficient on $log(1+HIndex)$ in column 1 is 0.074 (p-value < 0.001), more than twice as large as the comparable coefficient reported for overall selection in Table 2. The coefficient of 0.131 on the variable *HasHomerun* in column 2 indicates that a homerun publication increases the probability of progressing past the first stage by nearly 13 percentage points, or 36 percent relative to the probability for $HasHomerun = 0$ candidates of 0.36. 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 both specifications, and never significant. Based on $CommitteeTie$'s coefficient

remain similar in magnitude but are no longer statistically significant.

[24] In the final column of Appendix Table A1, we show that our main result on hometown committee ties is near-identical for this smaller sample where first stage data are available.

(-0.014) and its standard error (0.028) in column 1, we can rule out at the 95 percent level the existence of a positive effect of $CommitteeTie$ of greater than 4.1 percentage points (-0.014 + 0.028*1.96).

In columns 3 and 4, the dependent variable is *Elected*; we limit the sample to candidates who make it past the first selection stage. There is a very large effect of $CommitteeTie$ on second stage success across all specifications: a committee hometown tie is associated with a 15.8 percentage point increase in the probability of becoming a CAS/CAE fellow, conditional on making it through the first stage screening. Interestingly, the correlations between second stage success and our measures of research quality, $log(1 + HIndex)$ and $HasHomerun$, are much weaker - the coefficient on $log(1 + HIndex)$ in column 3 is close to zero, while the coefficient on $HasHomerun$ in column 4 is about 60 percent lower than its counterpart in column 2 (and significant only at the 10 percent level).

Overall, our results in Table 3 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).

3.2 Hometown ties and the quality of selected fellows

If hometown-connected fellows face a lower threshold for election, two further predictions follow: (a) the average quality of connected nominees will be lower; and (b) the quality of elected candidates (conditional on the pool of nominees) will be lower for connected candidates.

We explore these predictions in Table 4, where we report the results of the following specification:

$$Quality_{yi} = \alpha_{dy} + \beta_1 * CommitteeTie_{yi} + \beta_2 * NonCommitteeTie_{yi} + Controls_{yi} + \epsilon_{yi} \quad (2)$$

We do so for quality measures $log(1 + HIndex)$ and $HasHomerun$, examining their correlation with quality for the pool of candidates as they progress through the selection process. In columns 1 - 3 we employ $log(1+HIndex)$ as our outcome variable; the sample is comprised of all nominees in column 1, candidates progressing past the first stage in column 2, and elected candidates in column 3. We repeat these analyses in columns 4 - 6 using $HasHomerun$ as the outcome.

We find a small, statistically insignificant coefficient on $CommitteeTie$ in column 1, which includes the full sample of nominees. Thus, there is at best limited evidence of lower measurable quality for hometown-connected candidates in the nominee pool overall. The coefficient on $CommitteeTie$ increases as we move across the columns. In column 3, the pool of elected candidates, the coefficient on $CommitteeTie$, -0.392, is more than five times greater than in column 1, and significant at the 1 percent level. Comparing columns 2 and 3, it is clear that the negative selection of connected nominees (relative to unconnected ones) occurs primarily in the second (in-person) stage of selection.[25] Intriguingly, the coefficient on $NonCommitteeTie$ is *positive* and significant in the second stage. One natural interpretation is that this is a result of the directive for geographic diversity within the CAS/CAE, which we noted in Section 1 - if a hometown is already represented in a department, the quality bar may be higher for additional members.[26] In Appendix Table A3 we report results paralleling those in column 3, using alternative measures of research quality. In each case, $CommitteeTie$ is a negative predictor of elected fellow quality except when measured by Chinese H-Index.

[25]In Online Appendix Table OA6, we present results from a fixed effects Poisson (Quasi-ML) regression, with election-level clustering and also fixed effects. This analysis generates results that are very similar to those reported in Table 4, with comparable interpretation.

[26]However, note that in the second set of columns with $HasHomerun$ as our quality measure, the coefficient on $NonCommitteeTie$ is insignificant.

The size of this negative selection effect is large, and easy to see in the bar graph in Figure 2, which shows the median H-Index of candidates at each stage of selection. While the nominee pools for connected and unconnected candidates start out with comparable quality (median H-Index of 4 versus 4.5 respectively), among elected candidates the median H-Index of connected candidates is less than half that of unconnected ones (4.5 versus 10). Figure 2 also reveals a pattern that cannot be discerned from regression coefficients - while we observe positive selection on quality in the first stage for both groups, in the second stage there is negative selection overall (not just relative to unconnected candidates) for connected nominees who make it past the first selection stage. One might speculate that this reflects senior scholars' concerns about being displaced in the hometown *guanxi* network by more able - and ultimately more influential - fellows, in line with the idea that leaders face a tradeoff between loyalty and quality in choosing colleagues or advisors (see, for example, Egorov and Sonin (2011)).[27]

In the next three columns of Table 4, along with Figure 3, we repeat the preceding exercise with $HasHomerun$ as our quality measure. We obtain qualitatively very similar results, with a more intuitive interpretation. In particular, the coefficient on $CommitteeTie$ in the final column is -0.198. Given that the fraction of elected candidates with $CommitteeTie = 0$ who have a homerun publication is 0.398, it follows that a hometown tie cuts the probability that an elected fellow has a 100+ citation paper by half. Comparing the results of columns 1 and 3, it is clear that this effect comes almost entirely from the fellow election process rather than differences in the candidate pools.

In using these results to provide policy-relevant extrapolations, it is important to keep in mind that, while the impact of connections on individual quality is very large, in aggregate the effect of hometown ties needs to be scaled by their prevalence in the population. Recall that only 10 percent of all nominees - and 13.7 percent of elected fellows - are connected to

[27]This result also suggests no complementarity between connections and ability, and may imply – 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.

standing committee members via hometown ties. Thus, getting rid of hometown ties in the evaluation process would increase the homerun rate by only 2.7 percentage points for the membership overall (0.137*0.198). Of course, hometown ties represent only a single form of favoritism, so that the aggregate effect of all forms of favoritism may be much larger than the effect of hometown ties alone.

3.3 The consequences of CAS/CAE election for resource allocation

In our final set of analyses, we document the increase in influence and resources associated with CAS/CAE election, which complements our brief qualitative discussion in Section 2. We present here two pieces of empirical analysis. First, we document how membership affects an individual's chances of being appointed to a senior administrative post. Second, we show the relationship at the university-level between the number of fellows employed and government funding.

In our first set of analyses, we provide "event study" plots for the probability of receiving a university appointment of dean or president in the years around CAS/CAE election.

We divide the sample into candidates who are elected, and those who are nominated but never elected. For never-elected candidates who were nominated more than once, we focus on the first year in our sample when they receive a nomination as the event date. In practice, the patterns we observe are near-identical if we use their last year of nomination, or an average of all nomination dates.

In Figure 4, for each group we provide event plots showing the fraction of nominees that are appointed as dean or president of their institution in the $[-3, +3]$ year window around their nomination or election date. Since the most recent data for administrative appointments is from 2015, we use CAS/CAE nomination data for 2001 - 2011. The fraction of newly-elected CAS/CAE candidates obtaining appointments as dean or president, as shown by the solid line, increases markedly in the year of election and remains high for the

subsequent three years. By contrast, unsuccessful nominees exhibit no increase (perhaps even a small decrease) in the fraction receiving such appointments.[28]

A second channel of influence may come through funding. As we have observed previously, CAS/CAE fellowship tends to come late in a scholar's career. We are thus less interested in the funding that they access directly than in the funding they obtain for their collaborators or institutions.[29] While mapping out the personal and professional networks of candidates is beyond the scope of our study, we may examine how the presence of fellows affects university-level funding using publicly available data from China's Ministry of Education (MOE). These data, available in MOE yearbooks by institution, include total government grants and total scientists and researchers (including research staff) employed, for 2001-2013, excepting 2003-2004 when only municipal aggregates were available. The funding data include competitive grants (such as projects financed by China's National Science Foundation and Ministry of Science and Technology) as well as governmental budget allocations to each university. The latter part represents the vast majority of university-level funding, which is subject to considerable discretion on the part of MOE officials.[30]

Figure 5 provides a binned scatterplot showing the relationship between the number of fellows employed at each university and its total government funding. These scatterplots

[28]In Appendix Table A4 we show the effect of CAS/CAE election on senior administrative appointments in a regression framework. We focus on elected candidates, and include candidate and year fixed effects in all specifications (we also include third-order polynomial controls for age in most specifications given the strong (and non-monotonic) relationship between age and administrative appointments). Our results are roughly in line with those illustrated in Figure 4, with an estimated effect of CAS/CAE election on administrative appointments of 1.2 to 1.8 percentage points (significant at least at the 5 percent level in all cases). When we allow the effect of CAS/CAE fellowship to differ for connected versus unconnected candidates, the point estimate on *CommitteeTie*ElectedPost* is negative, but with a very large standard error.

[29]We also collected data on Chinese National Science Foundation funding for all the fellows in our sample, and conducted event study analyses paralleling those in Figure 4. We observe a modest increase in the probability of an individual receiving NSF funding in the year following CAS/CAE election, while we observe no such increase for unsuccessful nominees. Our regression estimates suggest that CAS/CAE election leads to a 2 percentage point increase in the probability of NSF funding (significant at the 5 percent level). However, the sums of money involved are relatively modest and award frequency quite rare, compared to the aggregate funding effects we document in the material that follows: NSF funding probability increases from about 10 percent to 12 percent, and the median award amount is 2.1 million RMB.

[30]Xu (2013) tells the story of CAE fellow and professor at Beijing Forestry University, Shen Guofang. According to the story, Shen wished to retire but his request was rejected because of the consequences - in terms of funding and prestige - for the university.

are residualized, removing university and year fixed effects, and present the data for all universities that employed at least one CAS/CAE fellow during 2001-2013. In addition, we control for total researchers, allowing its effect to differ across years owing to changes in the way that research staff are classified by the MOE across years. The scatterplot indicates a clear positive correlation between the number of fellows employed and total government funding. When we look at this relationship in a regression framework, again including university and year fixed effects, as well as controlling for the number of full-time researchers in each year, we estimate that a fellow is associated with an additional 63 million RMB in annual funding, or around US $9.5 million, significant at the 1 percent level.

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 39 percent more likely to be selected as CAS/CAE fellows, due entirely to higher success rates in the second (in-person) stage of the selection process. The hometown-connected candidates who gain election do so with considerably weaker scientific accomplishments than non-connected candidates - for example, elected candidates with hometown ties are about half as likely as unconnected candidates to have had a 100+ citation paper. Favoritism in selection into the CAS/CAE has potentially major effects on the allocation of research resources since, as we document, election increases the probability that scientists will receive high-level administrative appointments, and is associated with greater funding for the universities that employ them.

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

Acemoglu, D. (1995). Reward structures and the allocation of talent. *European Economic Review Vol. 39.1*, 17–33.

Altbach, P. G. (2012). *Paying the professoriate: A global comparison of compensation and contracts*. Routledge: Taylor & Francis Group.

Brandt, L., J. V. Biesebroeck, and Y. Zhang (2012). Creative accounting or creative destruction? firm-level productivity growth in chinese manufacturing. *Journal of Development Economics Vol. 97.2*, 339–351.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padr i Miquel (2015). The value of democracy: evidence from road building in kenya. *American Economic Review Vol. 105(6)*, 1817–1851.

Cohen, L. and C. J. Malloy (2014). Friends in high places. *American Economic Journal: Economic Policy Vol. 6(3)*, 63–91.

Crissman, L. W. (1967). The segmentary structure of urban overseas chiense communities. *Man, New Series Vol. 2, No. 2*, 185–204.

Dou, J. (1946). *Tongxiang Zuzhi Zhi Yanjiu: [Study on the hometown associations]*. Taipei: Taiwan Zhengzhong Publishing House.

Durante, R., G. Labartino, and R. Perotti (2014). Academic dynasties: Decentralization and familism in the italian academia. *Unpublished manuscript*.

Egorov, G. and K. Sonin (2011). Dictators and their viziers: Endogenizing the loyalty-competence trade-off. *Journal of the European Economic Association Vol. 9.5*, 903–930.

Freedman, M. (1960). Immigrants and associations: Chinese in nineteenth-century singapore. *Comparative Studies in Society and History Vol. 3(1)*, 25–48.

Hirsch, J. 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 Vol. 102.46*, 16569–16572.

Ho, P.-t. (1966). *Zhongguo Huiguan Shi Lun [Study on the history of Chinese guildhouses]*. Taipei: Taiwan Xuesheng Publishing House.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics Vol. 124.4*, 1403–1448.

Jia, R., M. Kudamatsu, and D. Seim (2015). Political selection in china: The complementary roles of connections and performance. *Journal of the European Economic Association Vol. 13(4)*, 631–688.

Kerri, J. N. (1976). Studying voluntary associations as adaptive mechanisms: A review of anthropological perspectives. *Current Anthropology Vol. 17(1)*, 23–47.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013). Trade liberalization and embedded institutional reform: Evidence from chinese exporters. *American Economic Review Vol. 103.6*, 2169–2195.

Lauren, C., A. Frazzini, and C. Malloy (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy Vol. 116(5)*, 951–979.

Lauren, C., A. Frazzini, and C. Malloy (2010). Sell-side school ties. *Journal of Finance Vol. 65(4)*, 1409–1437.

Li, D. (2015). Information, bias, and efficiency in expert evaluation: Evidence from the nih. *Unpublished manuscript*.

Li, L. (2011). Performing bribery in china: guanxi-practice, corruption with a human face. *Journal of Contemporary China Vol. 20.68*, 1–20.

Moll-Murata, C. (2008). Chinese guilds from the seventeenth to the twentieth centuries: An overview. *International Review of Social History Vol. 53*, 213–247.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics Vol. 106.2*, 503–30.

Ni, X. (2015). China's research and development spend. *Nature 30 April, VOL 520*.

Parsons, C. A., J. Sulaeman, and M. C. Yates (2011). Strike three: discrimination, incentives, and evaluation. *The American Economic Review Vol. 101, No. 4*, 1410–1435.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy Vol. 94, No. 5*, 1002–1037.

Shi, Y. and Y. Rao (2010). China's research culture. *Science Vol. 329, No. 5996*, 1128.

Siegel, J. (2007). Contingent political capital and international alliances: Evidence from south korea. *Administrative Science Quarterly Vol. 52, No. 4*, 621–666.

Xiao-Ping, C. and C. C. Chen (2004). On the intricacies of the chinese guanxi: A process model of guanxi development. *Asia Pacific Journal of Management Vol. 21*, 305–324.

Xu, Y. (2013). Hard for a fellow to retire? the issue of resource misallocation [yuanshi tuixiu nan beihou shi ziyuan de cuopei]. *Guangming Daily December 24*.

Young, A. (2003). Gold into base metals: Productivity growth in the people's republic of china during the reform period. *Journal of Political Economy Vol. 111.6*, 1220–1261.

Zhu, X. (2012). China's growth: Past, present and future. *Journal of Economic Perspectives Vol. 26(4)*, 103–124.

Zinovyeva, N. and M. Bagues (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics Vol. 7(2)*, 264–92.

Table 1a: Summary Statistics, Full Sample

Varible Name	Mean	StdDev	Observations
$CommitteeTie$	0.100	0.299	4921
$NonCommitteeTie$	0.332	0.471	4921
$Elected$	0.143	0.350	4921
$FirstStage$	0.404	0.491	4357
$Elected \vert FirstStage = 1$	0.338	0.473	1760
$log(1 + HIndex)$	1.677	1.271	4921
$Homeruns$	1.545	4.708	4921
$HasHomerun$	0.271	0.445	4921
$Doctorate$	0.457	0.498	4921
Age	58.393	8.846	4825
$PoliticallyConnected$	0.048	0.213	4921
$Dean$	0.403	0.491	4921
$Committee\_CollegeTie$	0.246	0.431	4921
$NonCommittee\_CollegeTie$	0.446	0.497	4921

Table 1b: Summary Statistics, by Committee Ties

	$CommitteeTie = 1$		$CommitteeTie = 0$		Difference	
Varible Name	Mean	StDev	Mean	StDev	Difference	t-statistic
$Elected$	0.196	0.397	0.137	0.344	0.059	3.557
$FirstStage$	0.388	0.488	0.406	0.491	-0.018	-0.718
$Elected \vert FirstStage = 1$	0.494	0.502	0.322	0.467	0.172	4.420
$log(1 + HIndex)$	1.617	1.238	1.683	1.274	-0.067	-1.104
$Homeruns$	1.418	3.945	1.559	4.785	-0.141	-0.629
$HasHomerun$	0.257	0.438	0.273	0.445	-0.016	-0.742
$Doctorate$	0.398	0.490	0.463	0.499	-0.065	-2.758
Age	59.713	8.474	58.247	8.875	1.465	3.448
$PoliticallyConnected$	0.035	0.183	0.049	0.216	-0.015	-1.429
$Dean$	0.373	0.484	0.406	0.491	-0.033	-1.403
$Committee\_CollegeTie$	0.402	0.491	0.229	0.420	0.173	8.506
$NonCommittee\_CollegeTie$	0.543	0.499	0.435	0.496	0.108	4.562

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

	$NonCommitteeTie = 1$		$NonCommitteeTie = 0$		Difference	
Varible Name	Mean	StDev	Mean	StDev	Difference	t-statistic
$Elected$	0.151	0.358	0.139	0.346	0.012	1.117
$FirstStage$	0.393	0.489	0.409	0.492	-0.016	-1.044
$Elected \vert FirstStage = 1$	0.368	0.483	0.324	0.468	0.045	1.860
$log(1 + HIndex)$	1.618	1.221	1.706	1.294	-0.088	-2.293
$Homeruns$	1.376	4.952	1.630	4.581	-0.254	-1.783
$HasHomerun$	0.245	0.430	0.284	0.451	-0.040	-2.949
$Doctorate$	0.390	0.488	0.490	0.500	-0.099	-6.617
Age	59.955	8.587	57.622	8.872	2.332	8.680
$PoliticallyConnected$	0.050	0.217	0.047	0.211	0.003	0.421
$Dean$	0.386	0.487	0.412	0.492	-0.026	-1.756
$Committee\_CollegeTie$	0.326	0.469	0.207	0.405	0.119	9.204
$NonCommittee\_CollegeTie$	0.533	0.499	0.402	0.490	0.131	8.743

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. $Homeruns$ is the number of homerun (100+ citations in English journals) publications by the year of nomination. $HasHomerun$ is an indicator variable denoting whether a candidate has at least one homerun publication by the year of nomination. Age is the candidate's age in the year of nomination. $PoliticallyConnected$ denotes candidates with a government rank of vice-Tingju (i.e., vice-mayor) 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 candidates 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	0.050***	0.053***	0.052***	0.048**	0.070***	0.055*	0.052**
	(0.019)	(0.020)	(0.020)	(0.024)	(0.025)	(0.032)	(0.026)
NonCommitteeTie	0.007	-0.000	0.001	-0.030*	-0.009	-0.011	0.006
	(0.011)	(0.011)	(0.011)	(0.017)	(0.014)	(0.021)	(0.013)
$log(1 + HIndex)$		0.030***		0.029***	0.037***	0.037***	0.027***
		(0.005)		(0.007)	(0.007)	(0.010)	(0.006)
Doctorate		0.020	0.030**	0.016	0.017	0.029	0.017
		(0.013)	(0.012)	(0.016)	(0.017)	(0.025)	(0.015)
Dean		0.008	0.012	0.007	0.000	0.021	0.004
		(0.011)	(0.011)	(0.013)	(0.014)	(0.021)	(0.012)
PoliticallyConnected		0.033	0.032	0.062**	0.040	0.139	0.027
		(0.024)	(0.024)	(0.029)	(0.032)	(0.137)	(0.024)
$log(Age)$		0.121***	0.141***	0.192***	0.231***	0.138**	0.119***
		(0.037)	(0.037)	(0.048)	(0.051)	(0.067)	(0.046)
Committee_CollegeTie		0.019	0.019	0.028*	-0.009	0.036	0.008
		(0.014)	(0.014)	(0.016)	(0.019)	(0.024)	(0.017)
NonCommittee_CollegeTie		0.009	0.008	0.016	-0.018	0.007	0.010
		(0.011)	(0.011)	(0.014)	(0.018)	(0.020)	(0.014)
HasHomerun			0.060***				
			(0.014)				
Department-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hometown FEs				Yes			
College FE					Yes		
Sample	Full	Full	Full	Full	Full	CAS	CAE
Observations	4921	4825	4825	4824	4641	1800	3025
R-Squared	.0235	.0335	.0316	.176	.152	.0334	.0222

Notes: Standard errors clustered by candidate in all regressions. The sample in columns 1 - 5 includes all candidates 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 nomination. *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 a government rank of vice-Tingju (i.e., vice-mayor) 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. *HasHomerun* is an indicator variable denoting whether a candidate has at least one publication with 100+ citations by the year of nomination. $log(1 + Hindex)$, *Doctorate*, 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 Each Selection Stage

	(1)	(2)	(3)	(4)
Dependent Variable	*FirstStage*		*Elected*	
CommitteeTie	-0.009	-0.014	0.158***	0.160***
	(0.028)	(0.028)	(0.043)	(0.043)
NonCommitteeTie	-0.019	-0.017	0.022	0.022
	(0.019)	(0.020)	(0.026)	(0.026)
$log(1 + HIndex)$	0.074***		0.009	
	(0.009)		(0.012)	
HasHomerun		0.131***		0.049*
		(0.025)		(0.028)
Department-Year FEs	Yes	Yes	Yes	Yes
Sample	Full	Full	FirstStage=1	FirstStage=1
Observations	4265	4265	1738	1738
R-Squared	.0696	.0602	.0517	.0529

Notes: Standard errors clustered by candidate in all regressions. The sample in columns 1 and 2 includes all candidates to the CAS during 2003-2011 and to the CAE during 2001-2013; the sample in columns 3 and 4 includes all candidates to the CAS during 2003-2011 and to the CAE during 2001-2013, who passed through the first stage of selection. The dependent variable in columns 1 and 2 is an indicator variable denoting whether candidate i made it through the first stage of candidate selection to the CAS/CAE in year y. The dependent variable in columns 3 and 4 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 nomination. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *HasHomerun* is an indicator variable denoting whether a candidate has at least one publication with 100+ citations by the year of nomination. $log(1+Hindex)$ is self-explanatory. Control variables include those in Table 2, with output suppressed to conserve space. See Online Appendix Tables OA4 and OA5 for full results, and see the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

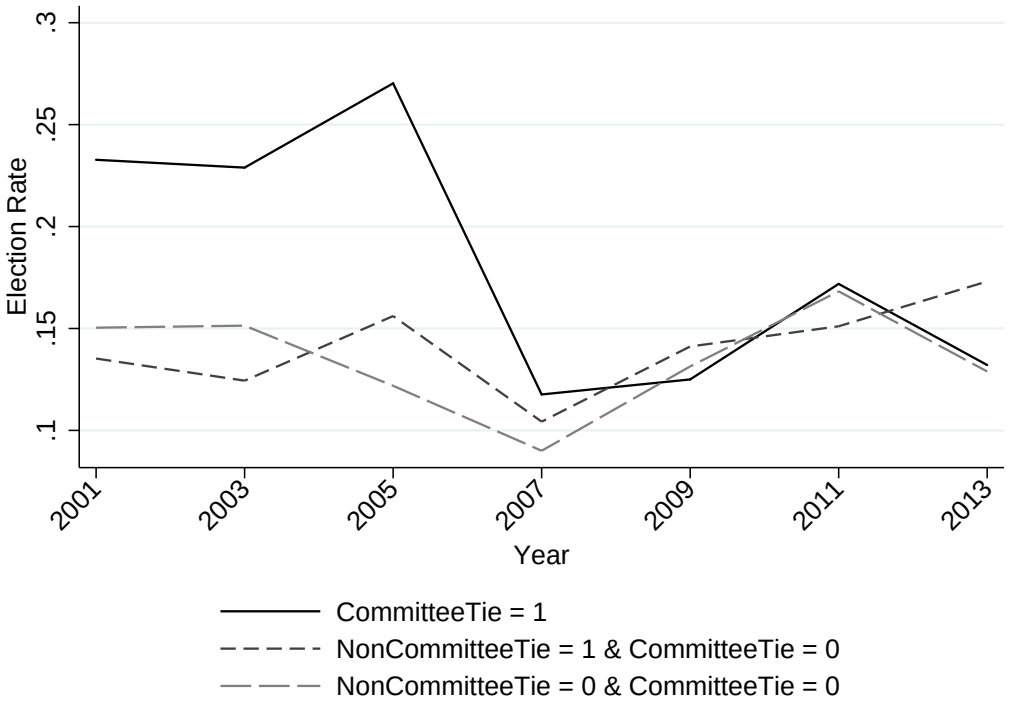
Table 4: Research Quality of Hometown-Connected versus Unconnected Candidates, at Different Stages of the Election Process

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable		$log(1+HIndex)$			$HasHomerun$	
$CommitteeTie$	-0.072	-0.110	-0.392***	-0.017	-0.054	-0.198***
	(0.056)	(0.091)	(0.117)	(0.024)	(0.039)	(0.050)
$NonCommitteeTie$	0.032	0.061	0.219**	-0.001	-0.001	0.059
	(0.040)	(0.065)	(0.086)	(0.016)	(0.026)	(0.038)
$Dean$	0.154***	0.213***	0.240***	0.010	0.015	0.003
	(0.038)	(0.059)	(0.083)	(0.015)	(0.024)	(0.036)
$PoliticallyConnected$	-0.049	-0.099	-0.087	-0.012	0.004	-0.075
	(0.091)	(0.144)	(0.175)	(0.032)	(0.057)	(0.078)
$log(Age)$	0.390***	0.432*	-0.025	-0.150**	-0.158	-0.293**
	(0.145)	(0.235)	(0.291)	(0.060)	(0.101)	(0.142)
$Doctorate$	0.545***	0.520***	0.485***	0.108***	0.109***	0.068
	(0.052)	(0.081)	(0.100)	(0.020)	(0.035)	(0.046)
$Committee\_CollegeTie$	0.018	0.019	0.008	0.004	-0.007	-0.022
	(0.044)	(0.068)	(0.092)	(0.016)	(0.027)	(0.041)
$NonCommittee\_CollegeTie$	0.008	0.018	-0.012	0.019	0.026	0.007
	(0.040)	(0.059)	(0.087)	(0.015)	(0.025)	(0.038)
Sample	Full	FirstStage=1	Elected=1	Full	FirstStage=1	Elected=1
Observations	4825	1738	700	4825	1738	700
R-Squared	.512	.537	.608	.353	.379	.418

Notes: Standard errors clustered by candidate in all regressions. All specifications include department-year fixed effects. The dependent variable in columns 1 - 3 is $log(1 + Hindex)$ while in columns 4 - 6 the dependent variable is $HasHomeRun$, an indicator variable denoting whether the candidate had at least one publication (100+ citations in English journals) at the time of nomination. Columns 3 and 6 include only candidates elected to the CAS/CAE. $NonCommitteeTie$ 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 a government rank of vice-Tingju (i.e., vice-mayor) 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. $HasHomerun$ is an indicator variable denoting whether a candidate has at least one publication with 100+ citations by the year of nomination. $log(1 + Hindex)$, $Doctorate$, and $log(Age)$ are self-explanatory. See the text for further details on variable construction.

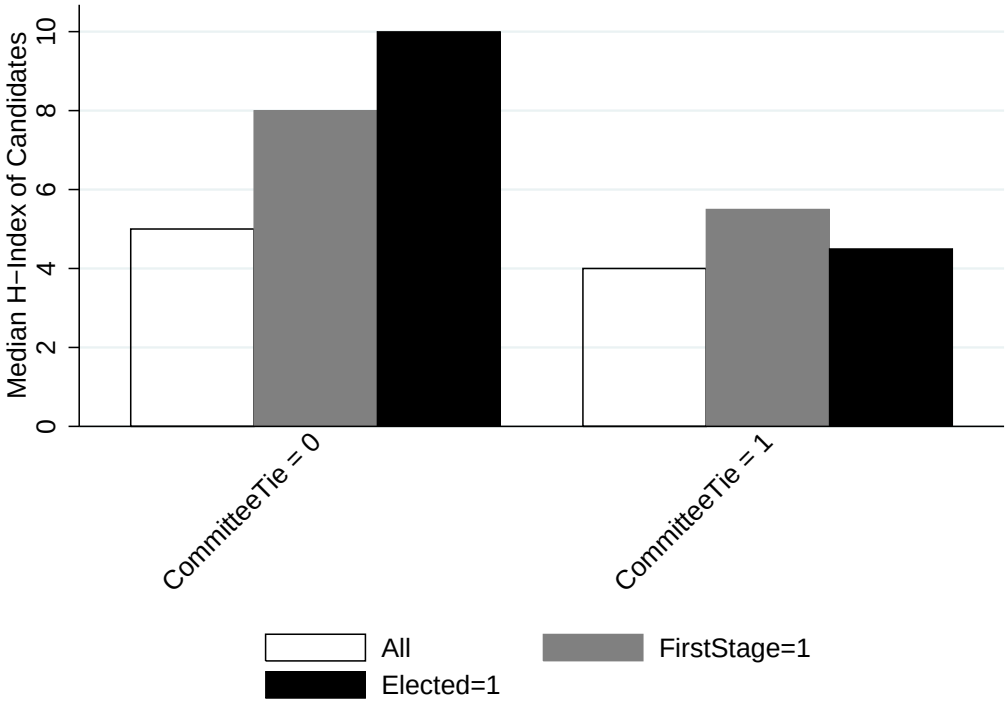
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Figure 1: Hometown Ties and Candidate Election Rates over Time



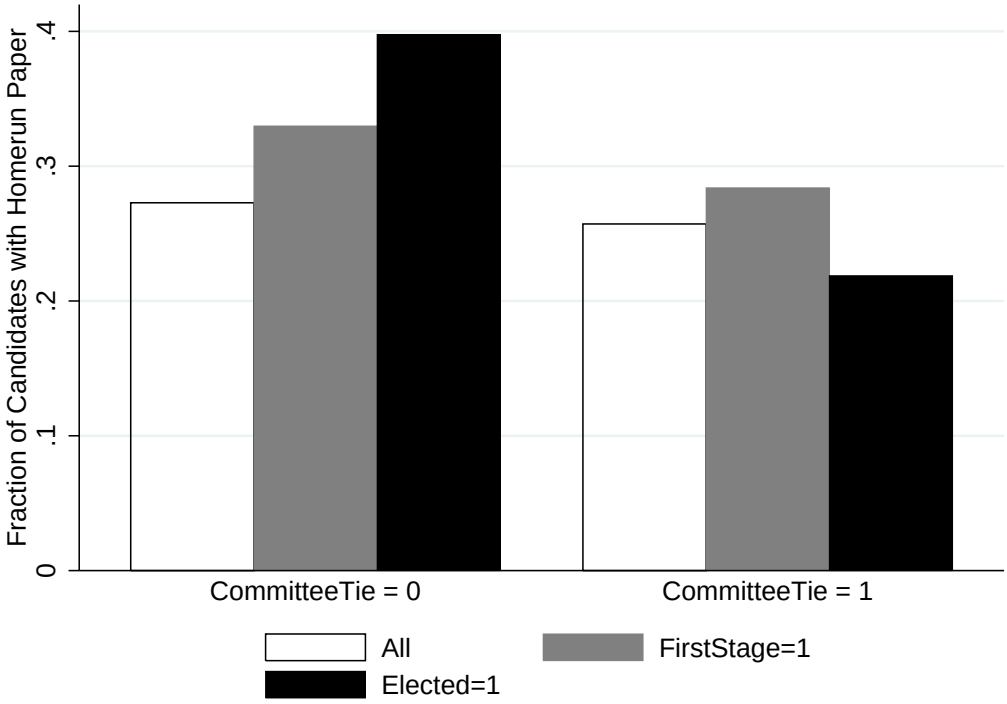
Notes: Each line provides the fraction of nominees elected to the CAS/CAE, disaggregated based on whether they have ties to fellows in the department of their nomination. The connections that characterize each group are provided in the figure legend.

Figure 2: Hometown Ties and Nominees' H-Indexes



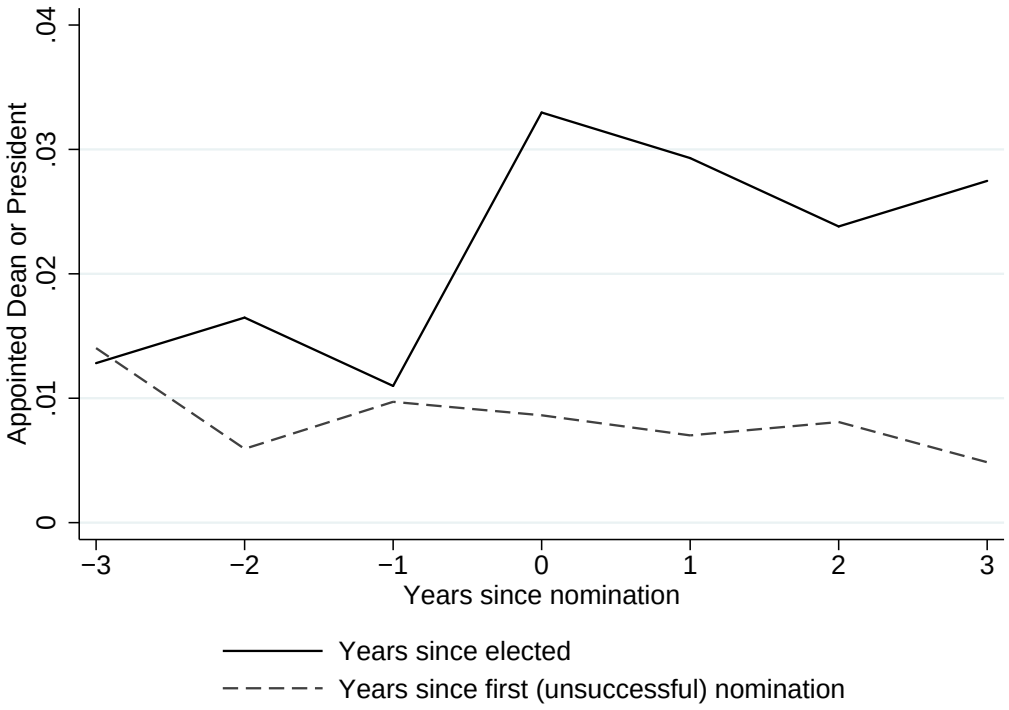
Notes: Each bar provides the median H-Index for a group of CAS/CAE candidates. The bars on the right are for the sample of candidates with hometown ties to standing committee members. The bars on the left are for candidates without such ties. The bars in each grouping are for progressively more selective samples of candidates. The white bars are for the full set of nominees. The gray bars are for nominees who progress past the first selection stage. The black bars are for candidates who are elected as fellows.

Figure 3: Hometown Ties and Nominees' Homerun Publications



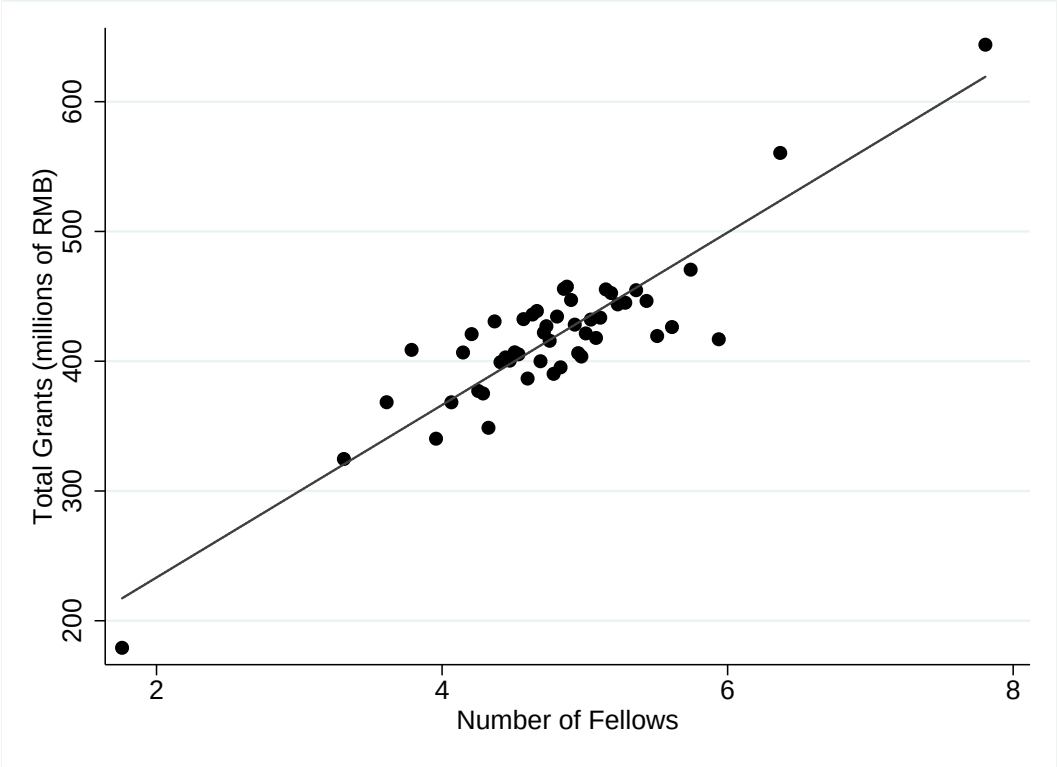
Notes: Each bar provides the fraction of nominees to the CAS/CAE that had at least one publication with 100+ citations at the time of nomination. The bars on the right are for the sample of candidates with hometown ties to standing committee members. The bars on the left are for candidates without such ties. The bars in each grouping are for progressively more selective samples of candidates. The white bars are for the full set of nominees. The gray bars are for nominees who progress past the first selection stage. The black bars are for candidates who are elected as fellows.

Figure 4: Election to the CAS/CAE and Appointment to Senior Administrative Posts



Notes: Each line provides the fraction of individuals in each group appointed as university dean or president around the years of nomination or appointment to the CAS/CAE. For both groups, $t = 0$ at the relevant year of nomination.

Figure 5: CAS/CAE Appointments and University Funding



Notes: This graph provides a binned scatterplot relating the number of CAS/CAE fellows at a university in a given year to its total grant funding. The specification used to generate the scatterplot includes fixed effects for university and year, and controls for the number of researchers in each year. The sample includes all universities with at least one fellow during the years 2001 - 2013, excluding 2003 and 2004 when no university-level data were available.

Table A1: Further Robustness Tests for Favoritism Results

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			*Elected*		
CommitteeTie	0.052***	0.052***	0.061**	0.094***	0.057***
	(0.020)	(0.020)	(0.025)	(0.027)	(0.022)
NonCommitteeTie	-0.001	0.001	-0.011	-0.003	0.001
	(0.012)	(0.011)	(0.015)	(0.016)	(0.012)
$log(1 + HIndex)$	0.030***	0.030***	0.033***	0.035***	0.028***
	(0.005)	(0.005)	(0.009)	(0.009)	(0.006)
Doctorate	0.020	0.015	0.010	0.001	0.020
	(0.013)	(0.013)	(0.019)	(0.011)	(0.013)
Dean	0.008	0.006	0.023	0.005	0.008
	(0.011)	(0.011)	(0.015)	(0.010)	(0.011)
PoliticallyConnected	0.033	0.033	0.051	0.033	0.033
	(0.024)	(0.024)	(0.039)	(0.023)	(0.024)
Committee_CollegeTie	0.019	0.020	0.013	0.020	0.016
	(0.014)	(0.014)	(0.018)	(0.014)	(0.015)
NonCommittee_CollegeTie	0.009	0.010	0.013	0.011	0.003
	(0.011)	(0.011)	(0.016)	(0.011)	(0.012)
CommitteeTie_Placebo	0.004				
	(0.012)				
Committee_EmployerTie			0.028		
			(0.022)		
NonCommittee_EmployerTie			-0.041**		
			(0.019)		
$CommitteeTie * I(Year \geq 2007)$				-0.096**	
				(0.039)	
$NonCommitteeTie * I(Year \geq 2007)$				0.015	
				(0.022)	
$log(1 + HIndex) * I(Year \geq 2007)$				-0.005	
				(0.011)	
Cohort FEs		Yes			
Employer FEs			Yes		
Sample	Full	Full	Full	Full	
Observations	4825	4825	4825	4921	4265
R-Squared	.0335	.0357	.191	.0335	.0328

Notes: Standard errors clustered by candidate in all regressions. All specifications include department-year fixed effects. The dependent variable in all columns is an indicator variable denoting whether candidate i was elected to the CAS/CAE in year y. Column 2 includes fixed effects for 5 year age cohorts, while column 3 includes fixed effects for a candidate's employer. The sample in column 5 excludes CAS nominees from 2013, to show robustness of our main results for the sample where data on first stage selection were unavailable. *CommitteeTie* is an indicator variable denoting that the candidate shared a 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. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with a government rank of vice-Tingju (i.e., vice-mayor) 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. *CommitteeTie_Placebo* denotes a hometown connection to a standing committee member *not* in the candidate's department. *Committee_EmployerTie* denotes a candidate who shares an employer with a standing committee member in his department of nomination. *NonCommittee_EmployerTie* denotes a candidate who shares an employer with a department fellow not on the standing committee. $I(Year \geq 2007)$ denotes observations from nomination years 2007 and later. $log(1 + Hindex)$, *Doctorate*, 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%.

Appendix Table A2: Robustness to Differing Controls for Research Quality

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			*Elected*		
$CommitteeTie$	0.053***	0.053***	0.051**	0.052***	0.052***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
$NonCommitteeTie$	-0.001	-0.000	0.001	0.000	0.001
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
$log(1 + Citations)$	0.012***				
	(0.002)				
$log(1 + Publications)$		0.017***			
		(0.004)			
$log(1 + ChineseHIndex)$			-0.007		
			(0.005)		
$Publications/1000$				0.428***	
				(0.091)	
H-Index Decile FEs					Yes
Observations	4825	4825	4825	4825	4825
R-Squared	.0331	.032	.0283	.0331	.0357

Notes: Standard errors clustered by candidate in all regressions. The sample includes all candidates to the CAS and CAE during 2001-2013, and all specifications include department-year fixed effects, as well as all controls in column 2 of Table 2. Column 5 additionally includes H-Index decile fixed effects (though with a larger fraction of the data for the bottom (H-Index=0) grouping). 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 nomination. $NonCommitteeTie$ denotes a hometown connection to a department fellow not on the standing committee. $Publications$, $ChineseHindex$ and $Citations$ are self-explanatory. $Publications$ are a candidate's total year-end publications in the year of nomination; $ChineseHindex$ and $Citations$ use all citations to these articles up to 2014. See the text for further details on variable construction. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table A3: Research Quality of Hometown-Connected versus Unconnected Candidates, Different Measures of Quality

	(1)	(2)	(3)	(4)
Dependent Variable	$log(1+Citations)$	$log(1+Publications)$	$log(1+Chinese_HIndex)$	$Publications/1000$
CommitteeTie	-0.928***	-0.508***	-0.127	-0.018*
	(0.280)	(0.181)	(0.136)	(0.010)
NonCommitteeTie	0.544**	0.327**	0.074	0.016*
	(0.212)	(0.131)	(0.094)	(0.009)
Dean	0.579***	0.517***	0.212**	0.011
	(0.203)	(0.127)	(0.096)	(0.008)
PoliticallyConnected	-0.117	-0.028	0.059	-0.008
	(0.406)	(0.261)	(0.202)	(0.016)
log(Age)	0.096	0.827*	1.031***	0.043
	(0.708)	(0.429)	(0.369)	(0.027)
Doctorate	1.014***	0.668***	0.303***	0.029***
	(0.242)	(0.158)	(0.105)	(0.011)
Committee_CollegeTie	0.059	0.003	0.086	0.001
	(0.219)	(0.141)	(0.097)	(0.007)
NonCommittee_CollegeTie	-0.065	0.001	0.026	0.003
	(0.213)	(0.130)	(0.092)	(0.008)
Observations	700	700	700	700
R-Squared	.601	.57	.319	.431

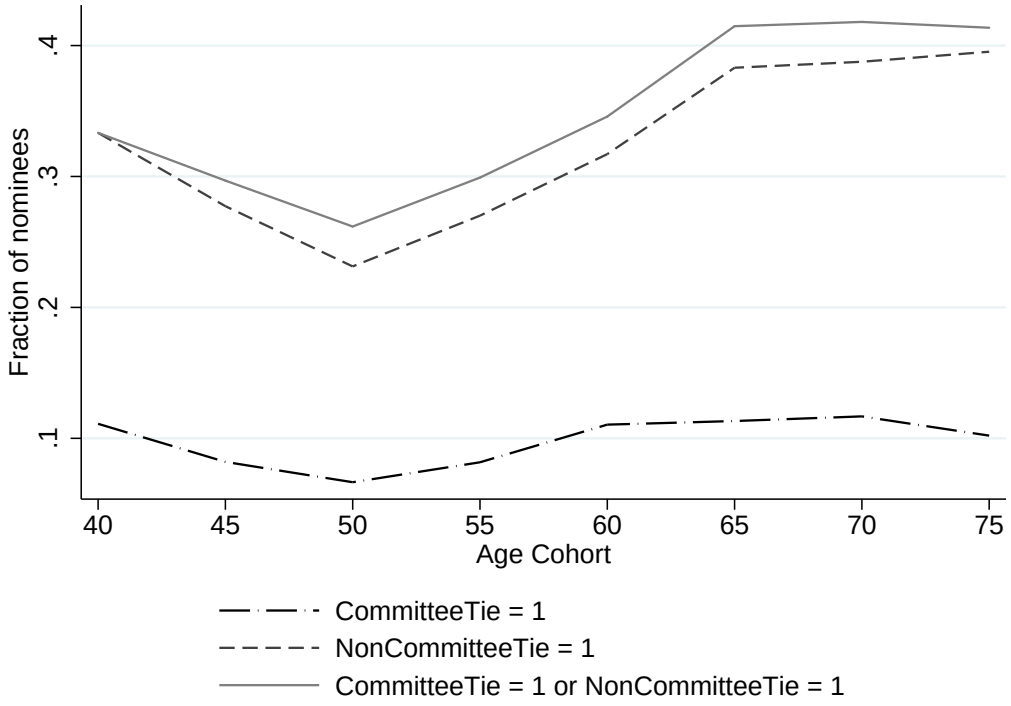
Notes: Standard errors clustered by candidate in all regressions. All specifications include department-year fixed effects. All specifications include only candidates elected to the CAS/CAE. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of nomination. *NonCommitteeTie* denotes a hometown connection to a department fellow not on the standing committee. *Dean* denotes a candidate holding an academic position of dean or higher, and *PoliticallyConnected* denotes candidates with a government rank of vice-Tingju (i.e., vice-mayor) 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. *ChineseHindex*, *Citations*, *Publications*, *Doctorate*, and *Age* are self-explanatory. *Publications* are a candidate's total year-end publications in the year of nomination; *ChineseHindex* and *Citations* use all citations to these articles up to 2014. See the text for further details on variable construction.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table A4: Hometown Ties and the Impact of CAS/CAE Membership on Senior Administrative Appointments

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			*LeaderAppointment*		
PostElection	0.012***	0.015***	0.017***	0.014***	0.018**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)
PostElection ∗ *CommitteeTie*			-0.007		
			(0.008)		
Age		0.205***	0.202***	0.185**	-0.197
		(0.071)	(0.071)	(0.072)	(0.497)
AgeSquared		-0.034***	-0.034***	-0.030**	0.042
		(0.012)	(0.012)	(0.012)	(0.087)
AgeCubed		0.002***	0.002***	0.002**	-0.003
		(0.001)	(0.001)	(0.001)	(0.005)
Candidate FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	11466	11376	11376	10831	4424
R-Squared	.0607	.0619	.062	.0663	.137

Notes: Standard errors clustered by candidate in all regressions. The dependent variable in all specifications is *LeaderAppointment*, an indicator variable denoting whether an individual was appointed president or dean in a given year. The sample in columns 1 - 3 includes all elected fellows from our main analysis in Table 2, and includes administrative appointments made during 1998 - 2015. Column 4 excludes election year observations to address concerns that election year administrative appointments may cause election to the CAS/CAE (i.e., reverse causation). Column 5 employs the sample used in Figure 4, limiting observations to the [-3,+3] window around election. *CommitteeTie* is an indicator variable denoting that the candidate shared a hometown with a standing committee member in the year of nomination. *PostElection* is an indicator variable denoting membership in the CAS/CAE (year of election is coded as $PostElection = 1$). *Age* is divided by ten to facilitate readability of coefficients.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix 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).