Political Pressure and the Direction of Research: Evidence from China's Academia

Daron Acemoglu
David Y. Yang
Jie Zhou*

July 1, 2021

— PRELIMINARY DRAFT —

Abstract

Freedom of inquiry is often considered as the foundation of innovation. Does political pressure impact the direction and quality of innovation in general, and academic research in particular? To answer this question, we collect comprehensive data on the scientific publications of researchers in the leading 109 Chinese universities and the leadership changes in these universities. We use NLP methods to measure the similarity between faculty members' and their leaders' research portfolios. We find that immediately after — and not before — the leaders take office, faculty members begin to shift their research direction towards that of their leaders. Such shifts are stronger among leaders with more political power and disciplines that have been historically more heavily targeted for academic persecution. Finally, faculty members' research publications receive lower citations after they pivot towards leaders' research. Taken together, our results suggest that political pressure impacts the direction of academic research at the expense of research quality.

*Acemoglu: Massachusetts Institute of Technology and NBER; email: daron@mit.edu. Yang: Harvard University and NBER; email: davidyang@fas.harvard.edu. Zhou: Massachusetts Institute of Technology; email: jiezhou@mit.edu. Junxi Liu, Michelle Han, Saier Wu, Peilin Yang, Andy Zhao, Hanyao Zhang, Xiaoyang Zhang, and Xinrui Zhou provided excellent research assistance.

1 Introduction

The freedom to experiment and pursue new and sometimes unconventional ideas is often viewed as fundamental to innovation (Hayek, 1978; Ridley, 2020). Conversely, there are ample examples of political interference from rulers or other powerful actors blocking innovations: ranging from the possibly apocryphal story of the Roman Emperor Tiberius killing an inventor who had come up with unbreakable glass and suppressing his innovation, to Queen Elizabeth I of England blocking William Lee's "stocking frame" knitting machine (Acemoglu and Robinson, 2012). Political inference in innovation does not typically take the form of explicitly blocking (let alone execution of innovators); but rather, political influence could often work by discouraging certain lines of inquiry and elevating political or other non-economic considerations ahead of innovation potential. This is no less true when it comes to academic research and innovation. Nevertheless, there is little direct evidence on whether political considerations impact the direction of innovation in general, and academic research in particular.

Academia in Mainland China provides an ideal setting for such an inquiry. Fostering innovation has been a central aim of the Chinese Communist Party, which recognizes the importance of technological progress for continued economic growth and has substantially increased funding for academic research during the past decades.[1] This desire notwithstanding, Chinese universities have enjoyed only limited autonomy, as exemplified by the fact that each faculty in every university has a centrally-appointed Communist Party representative in charge.[2] The juxtaposition of emphasis on innovation and the lack of academic freedom provides a unique opportunity to understand whether political pressures curtailing academic autonomy and freedom impact the direction and quality of innovation. Moreover, because China now accounts for a significant fraction of world research and innovation output, the potential distortion in Chinese academia is likely to have significant consequences for global innovation.

In this paper, we investigate this question by estimating the effects of new academic "leaders" (e.g., deans or department heads) on the type of research conducted by impacted faculty members across a large number of disciplines and Chinese universities. We are particularly interested in whether the politically-driven career concerns of faculty

[1] According to China's National Bureau of Statistics, fiscal spending on basic research rose from 25.8 billion RMB in 2013 to 42.3 billion RMB in 2018, a 64% increase; source: `http://www.stats.gov.cn/tjsj/ndsj/2019/indexeh.htm`.

[2] Academic freedom has further declined since President Xi Jinping's accession to power in November 2012. This is best illustrated by the decisions of many leading universities to remove clauses related to the "freedom of thought" from their charters; source: `https://reut.rs/39RVCNx`.

make them change their research so as to be more similar to their academic leaders' work and whether this impacts the quality of their research.

Our exploration of the linkages between political pressure and academic research is built on two new datasets. First, we collect information on the departmental organization across the social science disciplines in Chinese universities, as well as the departmental leadership transitions since 1990.[3] Second, we collect (close to) the universe of research publications by the faculty members and the department leaders during this period. We use natural language processing (NLP) methods to construct measures of research similarity between faculty members and their department leaders. We then investigate whether faculty members change their research style to make it more similar to that of their leaders.

Our baseline empirical strategy compares the similarity in research output between faculty members and their respective leaders, before and after the leaders take on the leadership position. Identification with this strategy relies on researcher-leader level variations over time, exploiting the fact that each researcher is observed under multiple leaders over the course of her academic career, and the leadership transitions take place at different times and with different frequencies across departments. This setup allows us to include a full set of researcher-leader pair fixed effects as well as calendar year fixed effects. In other words, we can account for other time-invariant factors that drive the similarity in research portfolios between researchers and their departmental leaders and changes thereof, and focus solely on changes in research similarity between faculty members and leaders driven by the appointment of the leaders.

We find that on average, the similarity in research output between faculty members and their respective leaders significantly and substantially increase after new leaders take office. Such changes indicate that the direction of research is shifted towards the portfolio of the current leader. The increase in research similarity between researchers and leaders emerge almost immediately after the leadership transition, and it persists for at least four years into the new leaders' tenure. This shift in research direction after leadership transition is robust to a range of alternative empirical specifications, including different NLP measures of research similarity. Bolstering confidence in our identification strategy, we do not find analogous changes in similarity *before* a leader takes up her position.

Shifts in research activities after leadership transition could be driven by two related but distinct channels. First, the appointment of a new leader might signal to researchers which research directions are favored by the Communist Party or funding bodies, in-

[3]We focus on social sciences for now, and we will expand to science and engineer disciplines subsequently.

ducing a change in research direction. Second, and more directly related to our focus, researchers may attempt to carry favor with the new leader who now have power over them. To distinguish the second mechanism — the impact of leaders on the research style of faculty under their direct supervision — we estimate the effects of new leader appointments in other institutions (and hence with no direct influence over researchers' careers) but in the same discipline (and hence, with the same potential signal value about what is being valued at higher levels). We also do this separately for new leaders in other top institutions (and still in the same discipline) where such signaling may be particularly salient. We do not find significant effect of leaders in the same discipline from other lower ranked universities, and a positive yet much smaller effect (relative to that of one's own leader) of leaders from higher ranked universities. This suggests that a moderate component of leadership transition effect could be attributed to signaling of research directions, but this does not account for the entirety of leadership transition in one's own department where political pressure is present. Importantly and reassuringly, we do not find any increase in research similarity between academic research in *other* disciplines and faculty members. These results thus increase our confidence that we are estimating the causal effect of new academic leaders on the research direction of faculty under their direct supervision.

Shifts in research activities after leadership transition could also be driven by standard career incentives unrelated to political pressures *per se*. To clarify the role of political factors, we show that the convergence of researchers' output towards the areas of their leaders is stronger when political interference is more powerful and more likely. First, compared to department chairs, we find that the department's Communist Party secretaries exert stronger influence on the faculty members' research direction. Tellingly, this is despite the fact that party secretaries are not leading academics. They are directly appointed by the Chinese Communist Party and represent closer alignment of the intention and direction of the Party and the Chinese authority. Second, we observe an increase in leaders' influence on faculty members' research directions after 2015, coinciding with the periods during which President Xi Jinping initiated a clampdown on (already limited) academic autonomy,[4] and many universities introduced "tenure track" reform, investing greater powers in universities on retention decisions.[5] The pattern that the similarity ef-

[4]President Xi called for Chinese authorities to exert greater control over the universities in 2016; source: `https://bit.ly/36aoXSf`.

[5]This in theory could have had ambiguous effects, because it simultaneously increased university autonomy and the power of departmental leaders in the careers of researchers. Albeit an extreme example, a recent murder case of a department Party Secretary in Fudan University due to his denial of tenure to a faculty member demonstrates the controls leaders can exert during the tenure promotion process (see `https://bit.ly/3qq3hKG` for details).

fects becomes intensified after 2015 suggests that the sum total of the effects of these two changes has been towards greater (political) career concerns for Chinese researchers.

We also explore whether academic persecution in the past may have persistently shaped academic environment and influenced the extent to which leaders are able to influence research directions of the faculty members. Following Wang and Kung (2021), we measure each academic discipline's likelihood of facing top-down targeting and selective persecution during the Cultural Revolution. We find that the more severely persecuted a discipline was four decades ago, the greater is the impact of leadership transition on research directions since 1990. Our interpretation is that the more severely-prosecuted disciplines are those where political career concerns still matter more, and thus our findings here suggest a specific mechanism for the legacy of past repression to have persisted to the present — via amplified political career concerns.

Finally, we document that the shifts in research direction due to leadership transition impose costs on quality of research. After the leadership transition, the research publications' citation counts among faculty members begin to decrease, particularly among papers that exhibit greatest similarity to the new leaders' research. Interestingly, the negative impact of leadership transition on faculty members' research quality enlarges as they experience multiple transitions, suggesting that switching research direction at relatively later stages of faculty members' career and switching multiple times are especially costly to maintaining quality of research. Moreover, we find that the increase in research similarity is significantly greater when new leaders have lower academic productivity, and stronger among department party secretaries who are relatively distant from the frontier academic research. This suggests that Chinese researchers are making their research more similar to precisely less productive and prolific leaders.

Taken together, these results suggest that political factors have a major impact on the direction and quality of academic research in China. Political pressure induces Chinese scholars to make their research more similar to their academic leaders, and our results suggest that this is driven by the influence of leaders who have direct supervisory power over them. In fact, we also document that departments begin to hire new faculty that are more similar to new leaders, suggesting a shift in the composition of the department as well. We interpret these findings as going beyond the normal career concerns and academic collaboration that might exist in any academic environment. We provide evidence that they are particularly pronounced for faculty members who have greater (political) career concerns themselves, and in departmental environment that may exhibit greater political pressure. Our results also indicate that quality of research suffers because of this political pressure — in other words, researchers under political influence may not have

realized their full potential had they align their research efforts with their comparative advantage.

Our paper is most closely related to the branch of the political economy literature investigating linkages between political factors and innovation. Much of the emphasis in this research has been on the risk of expropriation or political interference on entrepreneurship, investment and innovation, for example, as in (North et al., 2009; Acemoglu and Robinson, 2012). Potential future political threats from successful entrepreneurs may also encourage elites to block innovation to preserve their political power and rents (Acemoglu and Robinson, 2006).[6] Our mechanism is rather different as it shows the effects of local political pressure in academia — though the origin of this pinnacle pressure likely comes from national institutions. Our focus on academic research and innovation also connects our work to the growing literature on innovation economics, specifically, the various incentives that affect research production (e.g., Azoulay et al., 2011; Manso, 2011; Akcigit et al., 2018; Hill and Stein, 2021).

Our work also contributes to the literature on innovation in China. A large literature studies the Chinese economy and its growth in the past four decades. Recent works have carefully described the innovation landscape in China (e.g., Wei et al., 2017; Bombardini et al., 2018) and its potential implications for academic research (e.g., Freeman and Huang, 2015). More closely related to our paper is Jia et al. (2018), who document that academic leaders in economics departments in China's top universities tend to become more productive through co-authorship after they become leaders. This pattern — political power is used by academic leaders for their own benefit — suggests a different type of political distortion and is thus complementary to our research. We contribute to the understanding of how politics interact with innovation by providing, to the best of our knowledge, the first systematic analysis on the effects of political pressure on the direction and quality of academic research.

The rest of the paper is organized as follows. Section 2 describes the data sources used for this project; Section 3 describes the key measure of research similarity and the empirical strategy; Section 4 presents the results of leadership transition on career incentives; Section 5 presents evidence of politically motivated career concerns as an important source behind the baseline finding, and documents the costs on research quality; and finally, we conclude with discussions on the implications of the findings in Section 6.

[6] A nascent literature documents the alignment between the autocratic institutions and private innovation, particularly in the context of China. For example, Bai et al. (2020) examine how crony capitalism combined with local governments' competition can foster growth; Beraja et al. (2021) study how provision of government data and surveillance state's demand for AI can promote private innovation in the AI sector, due to the shareability of government data across multiple purposes.

2 Data

Our empirical analysis combines two primary datasets that we collect from scratch: *(i)* the structure of Chinese higher education institutions and the leadership information in the university departments; *(ii)* the scientific publications of all affiliates in these institutions; and *(iii)* number of citations of the publications. We now describe each of these datasets in turn.

2.1 University structure and departmental leadership

We first construct a dataset that describe the organizational structure and leadership of universities in Mainland China. We begin by examining all the social sciences departments among the top universities in China. We focus on the 109 universities that belong to the "Project 985" and "Project 211," two higher education ranking schemes that unambiguously list the top academic institutions in China.[7] Out of a total of 2,914 universities in Mainland China, these 109 top universities capture 70% of all research funding, and more than 50% of major scientific publications (Zhu, 2009; Zong and Zhang, 2019).

For each university, we collect the departmental organization structure for social sciences disciplines. We focus on "school" level bureaucratic structure, the organizational unit one level beneath the university in most Chinese universities, hence the leadership at the school level has most direct control over the personnel decisions of the faculty members. We standardize the organizational structure to make the school level definition comparable across universities.

We focus on schools and departments that are continuously active between 1990 and 2019, which is also the time window during which we collect publication records. For the schools and departments that cease to exist either due to splits or mergers, we track these changes and link schools and departments together so that past research activities in the previous academic units can be appropriately attributed to the corresponding departments today. This ensures that we don't have leader switches that only caused by changes in school structure. Overall, during the period of 1990 and 2019, there are on average 7.75 schools in a given university.

Broadly speaking, we group various schools and departments into a total of 11 disciplines: economics, management, business, finance; political science and public man-

[7] "Project 985" and "Project 211" are two major projects undertook by the Chinese government to promote the development and reputation of the Chinese higher education system by founding world-class universities. The universities included in these projects are top ranked in China, and many of them have since then ranked among the top 500 universities globally; source: `https://bit.ly/3ibF8Uo`.

agement; law; education; literature and media; history; psychology; philosophy, anthropology, ethnology, sociology; regional studies; foreign language; and Marxism. For the schools and departments that are interdisciplinary in nature, we classified them into the 11 disciplines by taking the disciplines within a school as children and the school as their parent. Then we group the parents into one classification if they have connected components. The details are described in Appendix A.

Finally, for each school, we identify its leaders during the 1990 to 2019 period from a variety of sources: university official websites, university yearbooks, *Baidu Baike* (a Chinese-language collaborative online encyclopedia), and various online reports that mention school leadership. We manually extract the department chairs and party secretaries for each school. For the years that we cannot locate precise leadership information, we employ several interpolation methods. [8] On average, each school experiences 2.8 leadership transitions during the three decades between 1990 and 2019. The average tenure of a given school chair is 5.8, though this varies fairly substantially across disciplines: ranging from 4.8 years in the discipline of Marxism, to 6.3 in the discipline of foreign language.

Note that similar to the bureaucratic structure in many organizations in China, universities and the schools within them have two parallel leadership posts: school chair in the academic track, and Chinese Communist Party secretary in the political track. We primarily focus on the leadership in the academic track since those individuals are scholars and have records of academic publications, making it relevant to study the potential re-pivoting of research effort by faculty members. In contrast, party secretaries often have no academic background and are rotated in from other Communist Party organs. We are able to identify a subset of school party secretaries who have academic track records, and we will compare their influences to the faculty members relative to that of the department chairs.

2.2 Research publications

We next construct a dataset of all scientific publications by scholars in the corresponding institutions during the three decades spanning 1990 and 2019. The scientific publication dataset serves two primary purposes: it provides the description of research output by

[8] Normally, if we find a faculty member showed up as a leader in the news in Year_1 and Year_2, we assume that this leader was holding this position from Year_1 to Year_2 if years in between are missing. When it is different leaders before and after the missing cell, if the missing years are no more than three, we conjecture that one of the leaders before and after the missing cell was still leader, and we interpolate by assigning the past leader to the missing cell. This is assuming that there may be less information about leaders that are about to step down, but for the new leader who just get on the position, it is more likely to have info about him/her. 16% of the missing department chairs are be solved under this assumption.

researchers which we rely on to construct our primarily outcomes of interest (describe in detail in Section 3.1). Moreover, the publication dataset allows us to retrospectively construct a roster of scholars affiliated with each institutions and the schools, since administrative records of complete school faculty roster are absent for most school and most years during the previous decades.

We rely on two major sources. First, *China National Knowledge Infrastructure (CNKI)*, a full-text database covering 90% of all official published Chinese journals. Second, *Wanfang Data*, which is a comprehensive database of Chinese journals, dissertations, and academic conferences. It provides access to 8,183 journals published in China and over 43.17 million articles, including 42.89 million full-text records (as of May 2019). To the extent that the coverage of these two databases do not fully overlap, they complement each other and when combined together, provide us with close to full coverage of scientific publications in Chinese journals.

For each researcher affiliated with the universities of interests described in Section 2.1, we collect all the papers she publishes during 1990 to 2019. We exclude publications on non-academic outlets such as newspaper opinion pieces. We also exclude publications as dissertation (e.g., part of the graduate studies) and other school internal journals. For each paper in the collection $\mathcal{D}$, we collect information about title, authors, publication year, abstract, and citations. This amounts to a total of 5,290,503 papers.

We then use the publication dataset to extract rosters of faculty members (and those who were or will become school and department leaders) in each academic unit. In a nutshell, we assign the academic affiliation to each author of a paper based on the publication information. Not all papers indicate affiliations at the school level; we thus assign the school level affiliation from any publication of a given author to all her papers. In order to rule out individuals who are affiliated with the school as student (and hence publishing sparsely) rather than faculty members, we use the dissertation database to locate the graduation year of a given researcher and count the post-graduation period as faculty affiliation. We also restrict the faculty members to those who publish more than 5 papers under a given affiliation and has publication records for more than 3 years, further excluding the ones that may publish with a temporary position such as visiting scholar. Our faculty roster extraction procedure performs well when we validate it with a set of contemporaneous school faculty list that we can obtain from the school's official website (see Appendix B for details).

This procedure provides us a list of faculty members affiliated with a particular school s, at a university u, in year t. Overall, we identify 42,395 active faculty members in social sciences disciplines in top universities during 1990 and 2019. On average, there are 62.2

faculty members in each school, ranging from 15.3 in the discipline of regional studies, to 103.0 in the discipline (category) of management, economics, finance, business. Each faculty member publishes on average 1.3 papers in a given year, ranging from 0.8 paper per year in the discipline of foreign language, to 2.4 papers per year in the discipline of psychology.

We notice a general trend of increased publication by scholars across all disciplines over the sampling period: the research productivity grows from 102 papers per year in 1991 to about 32,428 papers per year in 2018, reflecting the overall growth of Chinese academic institutions and research capacity over this period. We include year fixed effects in all baseline specifications to account for the secular trend in research activities.

2.3 Citation

In order to measure research quality, we also collect data on the count of citation of the research publications described above.

Data on citation counts is constructed separately from the publication dataset previously described. From the same CNKI and Wanfang databases, citation count data is collected and matched back to each paper in the publication dataset based on paper title and publication year. The maximum citation count is considered in instances where there are conflicting counts, such as between CNKI and Wanfang databases. This process yielded citation counts for 95.6% of the papers included in analysis. The remaining papers with missing citation counts are dropped from the dataset.

On average, each paper has a mean of 13.9 citations, with leaders averaging slightly more at 18.2 citations per paper, and faculty averaging slightly less at 13.7 per paper (see Appendix Table A.2 for summary statistics). Faculty at higher ranked universities also average higher citation counts: at top-ranked Peking University, the mean is almost double at 24 citations per paper.

3 Empirical strategy

In this section, we describe the empirical strategy, first on our core measurement of research similarity in Section 3.1, then on our baseline empirical specifications in Section 3.2.

3.1 Measurement of research similarity

3.1.1 Similarity of paper pairs

The building block of our measures of similarity in research outputs rests on constructing similarity scores over pairs of scientific publications. For each pair of papers in the paper collection $\mathcal{D}$, we construct a variety of measures of similarity $s : \mathcal{D} \times \mathcal{D} \mapsto \mathbb{R}^+$.

There are two broad classes of similarity measures: *(i)* non-parametric methods; and *(ii)* methods based on machine learning. Our baseline estimation uses the term frequency–inverse document frequency (TF-IDF), a non-parametric method, to measure similarity. We also use several alternative text-similarity measures, which we will describe in turn.[9]

Term frequency–inverse document frequency (TF-IDF) TF-IDF is statistical measure commonly used to evaluate how important a word is to a document in the context of a given corpus of documents (Biasi and Ma, 2020; Kelly et al., 2018). The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. Mathematically, for a word i, its TF-IDF score will be expressed as:

$$\text{TF-IDF}(i, d, \mathcal{D}) = \text{tf}(t, d) \times \text{idf}(t, \mathcal{D}), \tag{1}$$

where $\text{tf}(i, d)$ is the frequency of word i in document d, and $\text{idf}(i, \mathcal{D})$ is the log of one over the share of documents containing j in the corpus $\mathcal{D}$.

The collection of publications forms the text corpus $\mathcal{D}$ in which every individual document d is an abstract of a paper. By adopting the bag-of-words model, each document d can be represented as a vector $\mathbf{v}^d$ based on its words, discarding specific grammar and word order. The length of $\mathbf{v}^d$ is equal to the number of words in the vocabulary of the corpus $\mathcal{D}$[10].

Intuitively, we can let each element v_i^d be the number of times word i occurs in document d. Simply calculating the distance between the vectors of word frequencies to measure the similarity is problematic, however, because words that are common in every document (often called "stop words") will introduce bias in the similarity score.

With TF-IDF, we are able to map a document d to a vector $\mathbf{v}^d$ in which each element $v_i^d = \text{TF-IDF}(i, d, \mathcal{D})$. Then for two document $f, l \in \mathcal{D}$, the similarity measure is defined

[9] In this preliminary draft, we focus on results based on TF-IDF. We will introduce results using other similarity measures in subsequent draft.

[10] Here is an abuse of the notation $\mathcal{D}$. In previous context, $\mathcal{D}$ servers as the population of papers. But here we use it to refer a structured set of texts. Each text in this set is the abstract of a paper.

as:

$$s(f, l) = \mathbf{v}^f \cdot \mathbf{v}^l. \tag{2}$$

Doc2Vec Doc2Vec (Dai et al., 2015) is an unsupervised neural network algorithm which learns the fixed-length feature vectors from variable-length documents. Doc2Vec predicts each word in document d by utilizing the current document and a sliding window around the target word. We train a Doc2Vec model by using titles and abstracts of papers in paper collection $\mathcal{D}$ as the train set. After training, each document $d \in \mathcal{D}$ can be mapped to a document-unique feature vector $\mathbf{v}^d$, which intends to represent the concept of the document.

While training an informative neural network like Doc2Vec requires large train set and computation power, many researchers and institutions publicize their training results based on large corpus such as Wikipedia and newspapers. This allow us to use their results as pre-trained models and further finetune models to fit our paper collection $\mathcal{D}$. In our setting, we utilized two pre-trained Chinese word embeddings provided by Li et al. (2018): *(i)* Word2Vec trained with People's Daily News[11], and *(ii)* Word2Vec trained with Chinese Wikipedia. More details about the training process are in Appendix C.

After getting the vector representation of each document, similar to TF-IDF, we take cosine distance to measure the similarity between two document $f, l \in \mathcal{D}$: $s(f, l) = \mathbf{v}^f \cdot \mathbf{v}^l$.

3.1.2 Similarity score for a faculty-leader pair

Based on the similarity score between paper pairs, we can then construct measures of similarity in research portfolio for each pair of faculty member and her corresponding departmental leader.

Let $\mathcal{F}$ be the population of faculty and $\mathcal{L}$ be the set of all the leaders. For each faculty-leader pair $(F, L) \in \mathcal{F} \times \mathcal{L}$, we denote the set of papers published by the faculty member F in year t as $\mathcal{D}^F(t) = \{f_{t1}, f_{t2}, ..., f_{tn}\}$; the set of papers published by the leader L in year t as $\mathcal{D}^L(t) = \{l_{t1}, l_{t2}, ..., l_{tn}\}$; and the similarity score of a pair of papers (f, l) as $s(f, l)$.

Then for the faculty-leader pair, we construct the research similarity score at time t based on pairs of papers belonging in the following set: $\mathcal{P}^{(F,L)}(t) = \{(f, l) | f \in \mathcal{D}^F(t), l \in \cup_{k \leq t} \mathcal{D}^L(k)\}$. Namely, we compare all the papers published by the faculty in year t with all the papers that the leader has been published up until (and including) year t.

In order to capture the (potentially strategically) targeted pivoting of research activities specifically on a subset of salient papers, we define the similarity score between

[11] People's Daily is an official newspaper of the Chinese Communist Party and the largest newspaper group in China.

faculty-leader pair $i = (F, L)$ in year t as the maximum of the similarity among all pairs of papers published by these two researchers during the corresponding period: $y_{i,t} = \max\{s(f,l)|(f,l) \in \mathcal{P}^{(F,L)}(t)\}$.[12]

3.2 Empirical specification

Our baseline empirical strategy follows a modified event-study design. We compare the research similarity between pairs of faculty-leader, before and after the leaders take office, controlling for faculty-leader pair fixed effects and calendar time fixed effects. In other words, the shifts in faculty members' research portfolio is identified out of within faculty-leader pair variation over time. The baseline specification is:

$$Y_{i,t} = \sum_{l \neq -1; l=-3}^{4} \psi_l D_{i,t}^{l} + \alpha_i + \lambda_t + v_{i,t}, \tag{3}$$

where $Y_{i,t}$ is the similarity score for the faculty-leader pair i at time t; $D_{i,t}^{l}$ is an indicator for faculty-leader pair i being l periods away from initial treatment at calendar year t; α_i is a full set of the faculty-leader pair fixed effects; and λ_t is a full set of calendar time fixed effects. For each faculty-leader pair, we focus on the period of three years before and four years after the leadership transition. Our baseline results are robust to alternative choices of time window.

By conducting the analyses at faculty-leader pair level, we take advantage the fact that academic leadership transitions are not synchronized across universities. and departments. Such empirical strategy can account for any general shifts in research trends that could change researchers' effort and attention at a given time across all universities, as well as alleviating the typical concerns of selection into treatment. Moreover, to the extent that within a given discipline, departments across universities are not led by researchers working on similar topics and share particular research agendas, research direction shifts within faculty-leader pairs allow us to isolate department and leadership term specific effects that need not be identical throughout the departments within the discipline.

3.3 Threats to identification

The key identification assumption is that the only leader-faculty-pair-by-year level shocks that affect the similarity in research portfolio between the corresponding faculty member

[12] To capture the average shifts in research portfolio, we also define a specification of the research similarity based on the average similarity scores across all pairs of papers: $y_{i,t} = \frac{1}{|\mathcal{P}^{(F,L)}(t)|} \sum_{(f,l) \in \mathcal{P}^{(F,L)}(t)} s(f,l)$.

and her leader is the leadership transition that takes place in a particular year. We can examine the pre-trend leading up to the leadership transition to rule out any pending up effects leading to the leadership transition. We will also discuss a variety of placebo exercises that help rule out other leader-faculty-pair-by-year level shocks that could be correlated with the leadership transition.

One may be worried that leadership transition induces changes in faculty members' research productivity and research output quantity, thus changing the denominators of the research portfolio similarity between faculty and leaders. We examine faculty members' productivity changes leading up to and after leadership transition in the corresponding academic unit. Specifically, we regress the total number of academic publications per year on the time relative to leadership transition, controlling for faculty member fixed effects and calendar year fixed effects. This allows us to isolate the differential effects of productivity changes due to leadership transition, holding fixed the overall level of productivity of a particular faculty member and any time-specific general shock that affects researchers' productivity. In Appendix Figure A.3, we present the results. One can see that leadership transition does *not* induce changes in faculty members' research productivity (in terms of publication count), and there is no changes in productivity leading up to the leadership transition either.

4 Leadership transition and direction of research

4.1 Baseline results

We first examine the average effects of leadership transition on faculty members' research direction across all disciplines, all institutions, and over the past three decades. We pool all the sample together and estimate the baseline specification as described in Section 3.2. Our results are presented in Figure 1, where we plot the non-parametrically estimated ψ_l coefficients along with the corresponding 95% confidence intervals. The research portfolio similarity score between leaders and faculty members' publication at the year *prior to* leaders take office is normalized to zero, and the timing of leadership transition is marked by the vertical red line.

The estimates plotted in Figure 1 show a significant increase (by approximately 7%) in research similarity between faculty members and their leaders. There is no increase in similarity before the leader takes office, and the similarity index takes off immediately after leader turnover and persist for at least four years into the new leader's tenure. This timing, with no pre-trends, is reassuring concerning the validity of our identification

strategy. The absence of pre-trends also suggests that there are no anticipation effects before new leaders take office, or that researchers are selected to lead departments based on the similarity of the research portfolios with the rest of the faculty members.

Recall that our analysis analysis is conducted at the faculty-leader pair level. To the extent that faculty members' research activities and output within a department are diverse, the patterns depicted in Figure 1 suggest that, after the appointment of a new leader, faculty members pivot their heterogeneous research activities towards the same direction, getting closer to that of their leader's research portfolio. By the same token, the estimates also indicate that, after the appointment of a new leader, researchers pivot away from the research of past leaders.

A general concern for using TF-IDF is that this method may underestimate the similarity between two documents. The source of the underestimation comes from two assumptions underlying the TF-IDF method: *(i)* the meaning of the words are orthogonal to each other, and *(ii)* the word order doesn't matter when measuring the similarity between two texts. To relax the two assumptions, we calculate a vector $\mathbf{v}^d$ for each document d by training a Doc2Vec model (Dai et al., 2015). To further show that our results are robust to similarity scores measured under different contexts, we construct two different Doc2Vec similarity scores by combining our paper collection with Chinese Wikipedia and People's Daily respectively. Appendix Figure A.4 shows the impact of leader switch on the two types of Doc2Vec similarity scores. Consistent with the TF-IDF similarity score, before the new leader takes office, there is no rise in similarity, but once the new leader takes office, the similarity index keeps rising and lasts for at least four years.

We find that the effects concentrate in departments that are lower-ranked. In these departments, academic leaders tend to have substantial power over promotion and dismissal decisions, which is a reflection of the top-down nature of Chinese academic institutions and is potentially also related to the fact that many of the academic leaders are appointed explicitly or implicitly by the various administrative and bureaucratic units controlled by the Communist Party. We re-estimate our baseline specification on subsamples of the schools and departments, dividing them into three groups based on their ranking in the respective disciplines: top 10%, top 10%-40% and 40-70%. Note that even the relatively lowered ranked departments in our sample are among the top academic institutions in China as a result of our focus on the top 109 universities in the country. Appendix Figure A.5 presents the results. One can see that, indeed, while we observe increase in research similarity between faculty members and leaders after the leadership transition across all ranking groups, the effects are noticeably larger among the schools ranked below the 50th percentile and smallest for the top 10% group. Reassuringly, there

are no statistically significant pre-trends for any of the three groups.

Interestingly, the magnitude of the influence of leadership transition and the resulting shifts in political pressure on researchers' outputs differ noticeably across disciplines. We re-estimate the baseline specification separately by discipline and plot the similarity measure between leaders and faculty members at the five-year horizon in Appendix Figure A.6. The largest effects are in education, Marxism, management, economics, finance, business, law, media, philosophy, anthropology, ethnology, and sociology, while the effects are muted in political science, public management, and foreign language, and even negative, though imprecise, in history and psychology.[13] We will revisit the cross-discipline heterogeneity in the subsequent section to trace back its roots to the differential academic persecutions in the past.

4.2 Effects of leaders from other departments

Leaders may affect the research trajectories of faculty through two distinct but related channels: the local career concerns of faculty members under their direct jurisdiction, and the signals that all faculty in the discipline receive from the appointment of a leader with a particular research style and portfolio (which may also be related to career concerns since those heeding the signals may be more successful).

To isolate the effect working through the local career concerns, we estimate the baseline specification (3), including new leaders in the same discipline but other universities. Because signals from lower-ranked schools may be much less influential, we also separate leaders in the same discipline in other, higher-ranked As an additional placebo check, we also include leader switches in other disciplines, which should have no impact on the similarity between these new leaders in other disciplines and the faculty in question.

More specifically, to construct placebo leaders from the same disciplines but from other, higher-ranked departments ($\mathcal{L}2$), and from other, lower-ranked departments ($\mathcal{L}3$), we follow these steps.[14] First, for each $(F, L1, T1)$ in $\mathcal{L}1$, we filter all the leaders in our sample with the following criteria: *(i)* in other schools; *(ii)* in the same discipline; *(iii)* from departments that are top ranked (for $\mathcal{L}2$) or lower ranked ($\mathcal{L}3$);[15] and *(iv)* for whom

[13] We also find pre-trends in history and psychology, suggesting that these estimates are less reliable, perhaps because there are anticipatory changes in research styles and faculty members preemptively shift their research activities towards that of the incoming leaders. This would also be bolstered by the fact that leaders in these disciplines tend to have long tenure and potentially overlap in leadership during the formal leadership transition periods.

[14] We denote the actual leaders as Type I leaders, $\mathcal{L}1$.

[15] The ranking is based on the evaluation contacted by the Ministry of Education in 2019. The evaluation has 9 categories: A+: top 2%; A: 2%-5% (2% is not included); A-: 5%-10%; B+: 10%-20%; B: 20%-30%; B-: 30%-40; C+: 40%-50%; C: 50%-60%; C-: 60%-70%.

the leader switch happens in the range $[T1-3, T1+3]$. Second, in this filtered pool, we randomly select a leader $\mathcal{L}2$ (or $\mathcal{L}3$) and match her with the faculty member F. If there is no leader left after filtering, we skip the faculty-leader pair. Third, we put the leader back to the sample when we do the random matching for other observations in $\mathcal{L}1$. After matching each observation in $\mathcal{L}1$ with a new leader $\mathcal{L}2$ (or $\mathcal{L}3$), we can get the new sample $\mathcal{L}2$ (or $\mathcal{L}3$) in which each observation is $(F, L2, T2)$ (or $(F, L3, T3)$), where $T2$ (or $T3$) is the calendar year when $\mathcal{L}2$ (or $\mathcal{L}3$) becomes the leader of her own school.

We present the results for leaders in the same discipline but other departments in Figure 2, Panel A.

We do not find significant effect of leaders in the same discipline from other lower ranked universities, and a positive yet much smaller effect (relative to that of one's own leader) of leaders from higher ranked universities. This suggests that a moderate component of leadership transition effect could be attributed to signaling of research directions, but this does not account for the entirety of leadership transition in one's own department. In other words, the political power of leaders controlling the department influences faculty members' research directions, above and beyond the signals that one may infer about promising research areas from leaders' research.

4.3 Effects of leaders from other disciplines

Finally, we consider yet another type of leaders — those from different disciplines ($\mathcal{L}4$). These leaders would have no direct political influence on faculty members in different institutions, and they would not carry meaningful signals on research directions to scholars in different disciplines.

To construct these placebo leaders ($\mathcal{L}4$), the procedure is the same as that for $\mathcal{L}2$ and $\mathcal{L}3$, except for the conditions used for filtering leaders in the first step. Since and not being in the same discipline is the sufficient condition for being in different schools by construction, we only use one criterion to select the placebo leader $\mathcal{L}4$ here: $\mathcal{L}4$ is not in the same discipline as F. After matching each observation in $\mathcal{L}1$ with a new leader $\mathcal{L}4$ we can get the new sample $\mathcal{L}4$ in which each observation is $(F, L4, T4)$, where $T4$ is the calendar year when $\mathcal{L}4$ becomes the leader of her own school.

Figure 2, Panel B, turn to the results on the influence of academic leaders from other disciplines. Reassuringly, there is no effect of leadership transition among these leaders of other disciplines.

4.4 Political influence and hiring decisions

Having documented that political interference shifts the research direction of the faculty members towards their leaders', we next examine whether part of such change is driven by changes in the composition of faculty members. In particular, leaders are often very powerful, they may also shape the hiring direction of their department. To the extent that new hires share similar research direction with the leaders, this could contribute to the convergence of department's research direction towards that of the leaders.

To shed light on this question, we first examine whether leaders tend to hire faculty members whose research portfolio is closer to themselves. We focus on all faculty members who begin affiliation with a particular department or school during the 1990 to 2019 sampling period. We construct average research similarity scores between these faculty members and all leaders in the corresponding departments. The leaders who are in leadership position at the time faculty members' affiliation begins are the "hiring leader," under whose leadership (or potentially direct influence) the specific hiring decision is made. We investigate whether the research similarity between new hires and department leaders are particularly high if the leader is the hiring leader as well.

Table 1 presents results of the regression estimates of the faculty-hiring-leader pair's research similarity, relative of other faculty-leader pairs. We include a full set of faculty member fixed effects, which means that we are only exploiting variation across leaders. Our results indicate that the research similarity between faculty members and the hiring leaders are higher than that with other leaders in the same department. This indicate that leaders are more likely to hire new faculty members whose research portfolios are similar to the leaders themselves. Combined with the results presented so far, this pattern suggests that new leaders may be able to have very large impact on the research direction of the department by also influencing the identity of researchers.

The final question we investigate is whether, conditional on starting off as relatively closer to the hiring leader in terms of research direction, newly-hired faculty further shift the research portfolio to be closer to that of their leader. To examine this question, we re-estimate our baseline specification on sub-samples of the faculty-leader pairs of those who are existing faculty members, and of those who are new hires and their corresponding hiring leaders. Figure 3 presents the estimated results. While, as in our results so far, we can see the entire department shifting towards their leader's research portfolio, there is even a larger pivoting towards the leader's style among the newly-hired faculty. This pattern is consistent with politically-powerful leaders exerting a substantial effect on faculty who start their employment under their rule.

This impact on new hires and their research direction also suggests that the academic leadership can affect the composition of the department as well as its research trajectory, suggesting plausible misallocation costs resulting from political pressure in Chinese academia which we will examine in Section 5.3.

5 Politically-charged career incentives and their costs

There is politics and career concerns in every academic institution. Is academia in China different? In Section 5.1, we undertake several complementary exercises to show that the answer is likely yes. In Section 5.2, we examine the long term effects of past academic persecution on politically-charged career incentives. Finally, in Section 5.3, we assess the costs on research quality associated with politically-charged career incentives.

5.1 Academic leadership from the Communist Party

First, we examine whether the Communist Party secretaries of the departments exert stronger influence on the faculty members' direction of research. Similar to the bureaucratic structure of many organizations in China, two parallel leadership posts exist in universities and in each departments: chairs (or deans) and the Chinese Communist Party secretary. The party secretaries often have little academic background, and are directly appointed within the party organization and are rotated in from other Communist Party organs. We focus on a subset of school party secretaries who have academic track records — they are inactive scholars who embark on the political track within the academic leadership. We re-estimate our baseline specification, investigating the effects of leadership transition of the party secretaries on the research similarity between the faculty members within the department and the corresponding party secretaries. Figure 4, Panel A, presents the estimates. There are no statistically significant pre-trends prior to the party secretaries take office, and the research similarity between faculty members and their party secretaries begin to increase right afterwards. Albeit noisy estimated, we observe that the faculty members' research direction is, on average, three times more affected by the party secretaries, compared to that of the department chairs.

Next, we examine whether attempts to introduce tenure track system, aimed at converging Chinese academic institutions to the structure in the US and Europe and fostering greater academic performance, is effective at reducing the impact of the leaders that we identify. While several departments (such as the Department of Physics at Tsinghua University) introduced tenure track system in the early 2000s, vast majority of

the departments in the universities that we examine formally began the tenure track reform after 2014. We estimate our baseline specification separately before and after 2015. Note that this period also coincides with the early years of President Xi's term, during which political pressure in Chinese academia and restriction on academic freedom noticeably intensified; we cannot disentangle the tenure track system reform from the general shifts in academic environment. The results, plotted in Panel B in Figure 4, indicate that the increase in similarity with newly-appointed leaders is *greater* after 2015, suggesting that if anything, the introduction of tenure track reform amplifies political pressure and politically-charged career concerns experienced by the faculty members. This is consistent with ample anecdotal evidence that since the tenure track system requires explicit and formal personnel approval by the department leaders, it actually provided greater power to the leaders. In other words, the existing power structure within the academic institutions can be difficult to change, and attempts to reform could actually be captured by those that have power in the status quo.

Finally, we attempt to directly evaluate whether tenure track system reform in China imposing greater political pressure on faculty members. Then, for 20 social science departments at Peking University, we collect the starting date of tenure track system. If the tenure track reform starts before the faculty member enters the department, then the faculty member and all the corresponding faculty-leader pairs will be under the influence of the tenure track system reform[16]. With 5,770 faculty-leader pairs (among which 911 with tenure track), we separately estimate the effect of department head switch on similarity scores for faculty-leader pairs that are unaffected and affected by the tenure track system. The result, presented in Appendix Figure A.10, shows that with tenure track (grey dots), department heads have greater impact on their research similarity with faculty members. This suggests that more political pressures and politically sensitive career concerns are added by the tenure track system.

Taken together, these results suggest that the baseline career incentives are substantially exaggerated when they interact with forces of political interference and/or when there are no other guarantees of meritocratic promotion or advancement.

5.2 Persistence of past academic persecution

In fact, the extent to which leaders are able to influence research directions of the faculty members has institutional roots that are persistent. Specifically, we examine whether experiences of past academic persecution initiated by the Communist Party generates

[16]The tenure track system reform follows the provision that the old rule continues to apply to some existing situations while the new rule will apply to all future cases.

persistent increase in political pressure among contemporary Chinese academia. Following Wang and Kung (2021), we measure the academic disciplines' likelihood of facing top-down targeting and selective persecution during the Cultural Revolution. Wang and Kung (2021) show that academics in disciplines with higher ideological dissension, particularly those in Humanities and Social Sciences, faced disproportionately higher rates of intellectual persecution during the Cultural Revolution than their peers in Sciences, Applied Science and Engineering. To explore the influence of past academic persecution from the Communist party on political pressure among contemporary Chinese academia, we use the level of ideological dissension within a particular discipline to approximate for the amount of persecution faced during the Cultural Revolution.

We follow Wang and Kung (2021) in using the Discipline Classification and Code (*xueke fenlei yu daima*) developed by the National Standardization Management Committee (*guojia biaozhunhua guanli weiyuanhui*) to construct a measure of the variation in ideological dissension between academic disciplines. Under this classification scheme, each academic discipline is ranked based on the level of "consensus" or "paradigmatic development" of the discipline, with higher ranked disciplines corresponding with lower numerical values of classification codes. The extent to which consensus has been established within a discipline around general accepted theories, laws, frameworks, methods and beliefs serves as a proxy for the level of ideological dissension within that particular discipline. Higher consensus and thus lower dissension is associated with disciplines in Natural Sciences, corresponding to codes 110 - 190, while lower consensus and higher dissension is associated with Humanities and Social Sciences, corresponding to codes 710 - 910. The classification scheme assigns a unique code that identifies up to three levels of classification for each academic discipline: 62 "first-level" discipline groups (e.g., Economics) are divided into 676 "second-level" disciplines (e.g., Labor Economics), and then further subdivided into 2,382 "third-level" disciplines (e.g., Labor Economic History). As in Wang and Kung (2021), we utilize the second-level classification code for analysis. A faculty member in the discipline of "Economics" would receive a unique second-level classification code based on whether they specialized in political economy, labor economics, development economics, business economics, economic history, and so forth. Based on the ranking of second-level classification codes, we construct a normalized index which assigns zero to the discipline with the lowest level of ideological dissension and one to the discipline with the highest. Thus, each faculty member is assigned an index capturing the rank of ideological dissension in the field they are in, and subsequent level of exposure to selective persecution during the Cultural Revolution.

We re-estimate the baseline specification on the effects of leadership transition on fac-

ulty members' research direction, as described in Section 3.2, discipline by discipline. In Figure 5, we plot the estimated leadership transition effects against the rank of ideological dissension (indicating stronger intellectual persecution during the Cultural Revolution) for the corresponding discipline. One observes a significantly positive relationship: namely, the more severely persecuted a discipline was four decades ago, the greater is the impact of leadership transition on research directions since 1990, suggesting a legacy of political pressure and deprivation of academic freedom that was stemmed from the Cultural Revolution.

5.3 Costs on research quality

Do these politically-charged incentives impost costs on research quality? We deploy three strategies to answer this question.

First, we examine whether the leadership transition affects the number of citations of the papers published by the faculty members, which is a common proxy for research quality. Specifically, we re-estimate the baseline specification as described in Section 3.2, but examining the average citation as outcome of interest instead. The results are plotted in Figure 6. We find that immediately after the leadership transition (and not before), the citation counts of faculty members' research publications that are most similar to their new leaders among faculty members begin to decrease. Interestingly, this is not the case for faculty members' research publications that are least similar to their new leaders. This suggests that the pivot of research direction towards the new leaders move faculty members away from their academic strength and come at a cost of producing lower quality research on average.

Second, we examine which forms of leadership transition induce the strongest tow on research quality. In Appendix Figure A.7, Panel A, we separately estimate leadership transition's impact on faculty members' research publication citation, depending on how many times have the faculty members experienced leadership transition since joining the department. We find that while the 1st leadership transition that faculty members go through induce relatively muted reduction on their research quality, the negative impact increases in the 2nd transition and particularly those beyond the 2nd one. This suggests that switching research direction at relatively later stages of faculty members' career and switching multiple times are especially costly to maintaining quality of research. In Appendix Figure A.7, Panel B, we separately estimate leadership transition's impact on faculty members' research publication citation, depending on whether the new leaders conduct relatively high or low quality research prior to them becoming leaders. Note

that leaders are in general leading scholars within the departments, so even the relatively lower quality leaders are still on average conducting research that generate more cites that the rest of the faculty members. We observe that switching to a leader who had a track record of conducting (relatively) lower quality research has an especially large negative effect on the research quality among faculty members.

Third, we assess the changes in relationship between faculty-leader's research similarity and faculty members' research citations over the courses of the leaders' term. Appendix Figure A.8 plots this relationship estimated at the faculty-leader pair level for each year before and after corresponding leaders take office. We find that similarity with future leaders' research is associated with significant higher number of citations prior to the leaders take official post. This is consistent with the fact that faculty members could naturally sort into future leaders' areas of research based on their comparative advantage, and to the extent that (future) leaders may be more likely to work on important areas of research, faculty members' own research may receive more citations as well. However, the citations premium of conducting research similar with the leaders' sharply decreases immediately after the leaders take office. This suggests that after leadership transition, such sorting becomes diluted due to political pressure and other career incentives, and as a result, the citation premium diminishes.

Finally, we investigate whether the role of academic leaders' on faculty members' direction of research are stronger among leaders who are less academically oriented. To operationalize this, we re-estimate the baseline specifications on sub-samples of faculty-leader pairs divided based on the leaders' relative productivity within the profession. The academically less productive leaders are further from the academic track, and closer to the political track in terms of aiming for administrative and political leadership within the academic institutions. Specifically, in Appendix Figure A.9, the yellow (grey) dots represent results estimated from faculty-leader pairs of whom the leaders publish more (less) papers prior to them becoming leaders than the median faculty members in the discipline. We observe that the influence that leaders exert on the faculty members under their jurisdiction is substantially stronger among less productive, less academically oriented leaders (though the more productive leaders also exhibit positive effects). At face value, this finding suggests that misallocation costs could be significant, since the increase in similarity is driven by academics trying to adopt the styles of academically less inclined and successful leaders. This pattern is also corroborated with the result that researchers are differentially more affected by the party secretaries, who are leaders less likely to be on the academic track and hence further away on the research frontier.

6 Conclusion

Throughout history, many authoritarian regimes have been suspicious of innovation, research and new technologies, and have often discouraged or even sometimes blocked them (Mokyr, 1992; Acemoglu and Robinson, 2012). Even Soviet Russia, which poured huge resources into military and nuclear technologies and cultivated top-quality research in chemistry, physics and mathematics, was opposed to new technologies that were deemed to be destabilizing (Fitzpatrick, 1999). Modern-day China may thus be viewed as an almost unique case of an authoritarian regime deeply committed to innovation. But do the authoritarian political system and its reverberations throughout Chinese bureaucracy and society still distort the direction of research and suppress its quality? This question is central not just for the future of China's growth, but also for global innovation, especially given China's growing role therein. Nevertheless, we are not aware of any systematic investigation of the impact of political factors in the direction and quality of research and innovation.

In this paper, we undertake such a study, focusing on the top 109 Chinese universities. We exploit the appointment of new departmental leaders, who typically have extensive powers, partly as a reflection of the Chinese authoritarian system and how Chinese academia has been organized. The main question we explore is whether the appointment of new leaders leads to a change in the research portfolio and style of faculty members under their jurisdiction. We build a data set comprising the academic publications of all leaders and faculty members in these universities, and using NLP methods we construct measures of similarity between leaders' and faculty members' research output.

Our main finding is a strong increase in research similarity between a leader and the faculty under her jurisdiction. Reassuringly, there is no pre-trend — the increase in similarity starts after the new leader takes up office. We also show that leaders in the same discipline but in other universities as well as leaders in other disciplines do not have similar effects. We interpret these results as being due to politically-charged career concerns in Chinese academia that primarily impact faculty via local pressures exerted by (or implicitly felt from) the leaders under whose jurisdiction they are.

Career concerns are not confined to Chinese academia or authoritarian settings, however. It is plausible to presume that analogous changes in research strategy may happen in academic systems with greater autonomy. To build the case that the patterns we describe go beyond what would happen in situations where there is greater autonomy, less political interference and better institutional safeguards for meritocratic promotions and external review, we adopt a number of complementary strategies. First, the effects of new

leaders on their faculty is more pronounced in lower-ranked departments, which typically lack procedures for external review. Second, we document that they have grown in importance after 2015, when academic autonomy started declining further under President Xi. Third, we show that they are significantly larger for faculty who themselves have political career concerns — in particular, those who later become leaders themselves.

Do political pressures affect the quality as well as the direction of research? We provide, albeit indirect, evidence that the answer is likely yes. First, the increase in research similarity between leaders and faculty under their jurisdiction is much higher for leaders who themselves have lower academic productivity, which suggests that faculty are emulating the research agenda of low-productivity leaders. Second, we show that leaders hire new faculty that have a similar research portfolio to theirs, and these newly-hired faculty further change their research trajectory to make their output even more similar to that of the leader. Finally, even though we find no impact on the total number of papers, we are currently exploring whether citation-weighted output of faculty is impacted while they are trying to make their output more similar to their leaders'.

We view our paper as a first step in a research agenda that explores the relationship between the direction and quality of innovation and political factors, originating both from national institutions and local organizations. This agenda appears important for several reasons. As new technologies such as AI, nanotechnology and new materials become increasingly important, the role of research and innovation for global prosperity is likely to grow. However, how different political systems and local and global incentives, coming from political or other considerations, impact the direction of this research is unclear. This is critical for the future of Chinese growth, which can be seen as a unique historical experiment in combining an authoritarian political system with a relentless focus on innovation. It is also central for understanding the forces impacting the nature of academic research anywhere and designing better academic institutions under democratic institutions as well as contributing to academic freedom and autonomy in less democratic environments.

In this light, there are several interesting research areas that can be further explored. First, it is important to conduct similar studies in other contexts, to benefit from a comparative perspective — in particular to see whether similar politically-charged career concerns are present in the academia of less authoritarian countries. Second, in other, more data-rich environments, it may be possible to look at other characteristics of leaders, for example, where they have obtained their degree and how they have risen in the academic hierarchy. Last but not least, a similar analysis in the context of corporate innovation, for example, linking the nature of patents to the priorities and organizational structure of the

firms under which the research is being conducted, would be a very fruitful area.

References

Acemoglu, Daron and James A Robinson, "Economic Backwardness in Political Perspective," *American Political Science Review*, February 2006, *100* (1), 1–17.

_ **and** _ , *Why Nations Fail* The Origins of Power, Prosperity, and Poverty, New York: Crown Business, August 2012.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi, "Dancing with the Stars: Innovation Through Interactions," *NBER Working Paper*, March 2018, pp. 1–62.

Azoulay, Pierre, Joshua S Graff Zivin, and Gustavo Manso, "Incentives and creativity: evidence from the academic life sciences," *The RAND Journal of Economics*, September 2011, *42* (3), 527–554.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Song, "Special Deals with Chinese Characteristics," *NBER Macroeconomics Annual*, January 2020, *34*, 341–379.

Beraja, Martin, David Y. Yang, and Noam Yuchtman, "Data-intensive Innovation and the State: Evidence from AI Firms in China," *NBER Working Paper*, January 2021, pp. 1–84.

Biasi, Barbara and Song Ma, "The Education-Innovation Gap," Technical Report, working paper, Yale 2020.

Bombardini, Matilde, Bingjing Li, and Ruoying Wang, "Import Competition and Innovation: Evidence from China," *Working Paper*, January 2018, pp. 1–44.

Dai, Andrew M, Christopher Olah, and Quoc V Le, "Document embedding with paragraph vectors," *arXiv preprint arXiv:1507.07998*, 2015.

Fitzpatrick, Sheila, *Everyday Stalinism* Ordinary Life in Extraordinary Times: Soviet Russia in the 1930s, Oxford University Press, March 1999.

Freeman, Richard B and Wei Huang, "China's "Great Leap Forward" in Science and Engineering," *NBER Working Paper*, April 2015, pp. 1–31.

Hayek, Friedrich August, *The Constitution of Liberty*, University of Chicago Press, 1978.

Hill, Ryan and Carolyn Stein, "Race to the Bottom: Competition and Quality in Science," *Working Paper*, 2021.

Jia, Ruixue, Huihua Nie, and Wei Xiao, "Power and Productivity in Chinese Academia," *Working Paper*, December 2018, pp. 1–37.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," Technical Report, National Bureau of Economic Research 2018.

Li, Shen, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, and Xiaoyong Du, "Analogical Reasoning on Chinese Morphological and Semantic Relations," in "Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)" Association for Computational Linguistics 2018, pp. 138–143.

Manso, Gustavo, "Motivating Innovation," *The Journal of Finance*, September 2011, *66* (5), 1823–1860.

Mation, Lucas Ferreira and Aguinaldo Maciente, "GROUP_TWOWAY: Stata module to group observations by the connected components of two variables," 2014.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

Mokyr, Joel, *The Lever of Riches* Technological Creativity and Economic Progress, Oxford University Press, April 1992.

North, Douglass C, John Joseph Wallis, and Barry R Weingast, *Violence and Social Orders* A Conceptual Framework for Interpreting Recorded Human History, Cambridge: Cambridge University Press, February 2009.

Ridley, Matt, *How Innovation Works* And Why It Flourishes in Freedom, HarperCollins, May 2020.

Wang, Alina Yue and James Kai-Sing Kung, "Silence is Golden: Ideological Dissension and Selective Persecution in China's Cultural Revolution," *Working Paper*, 03 2021, pp. 1 – 57.

Wei, Shang-Jin, Zhuan Xie, and Xiaobo Zhang, "From "Made in China" to "Innovated in China": Necessity, Prospect, and Challenges," *Journal of Economic Perspectives*, February 2017, *31* (1), 49–70.

Zhu, Junwen, " SCIE : 1978-2007." PhD dissertation, : 2009.

Zong, Xiaohua and Wei Zhang, "Establishing world-class universities in China: deploying a quasi-experimental design to evaluate the net effects of Project 985," *Studies in Higher Education*, 2019, *44* (3), 417–431.

Figures and Tables

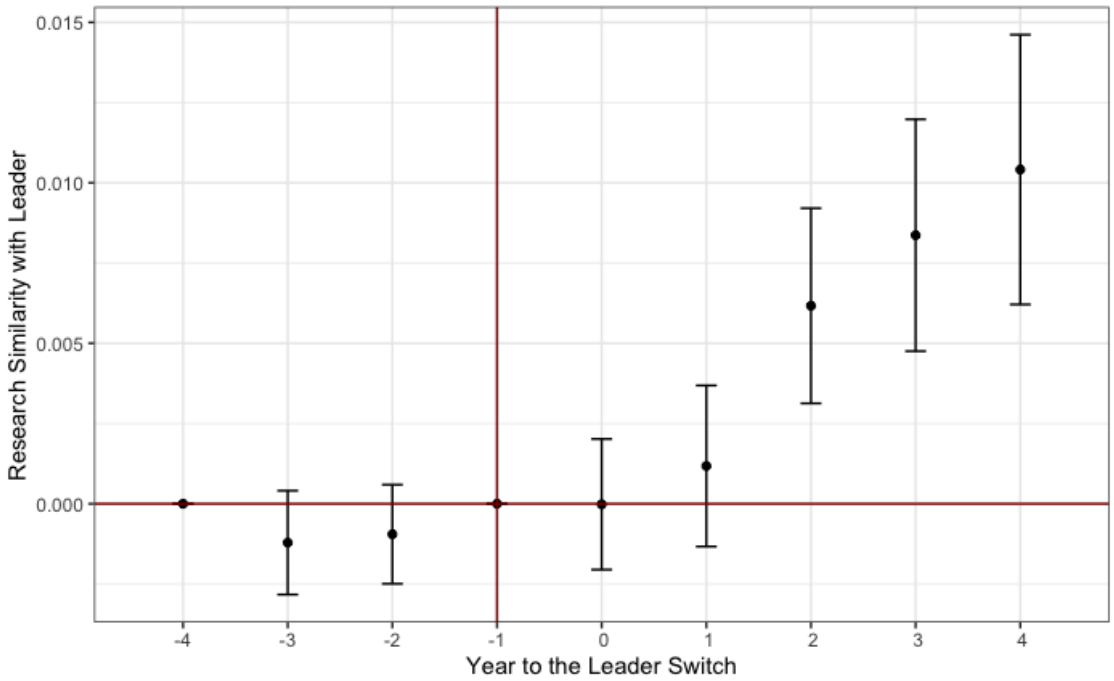
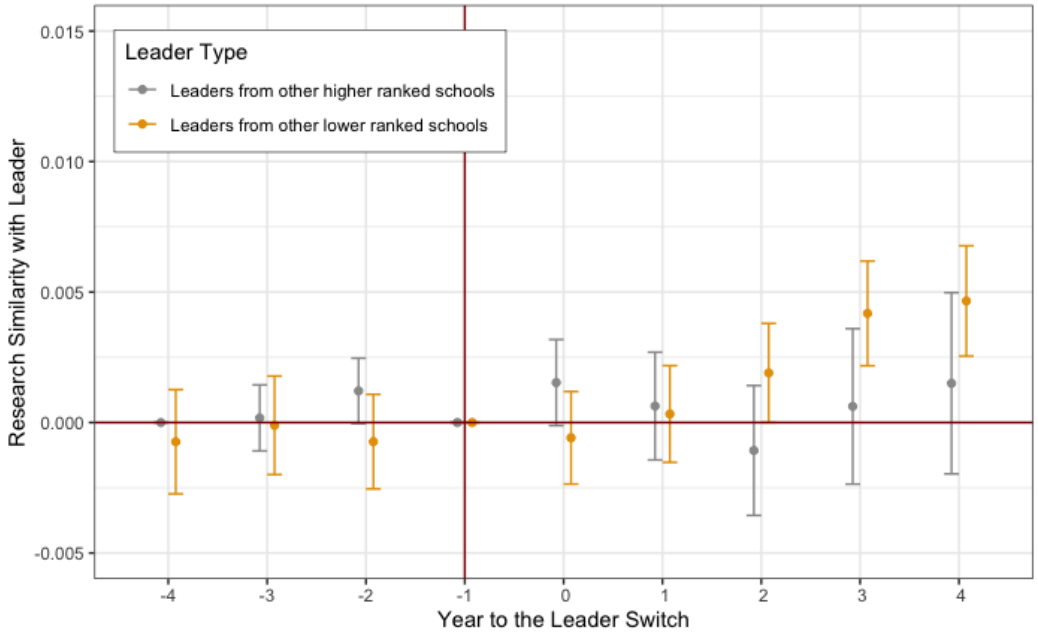
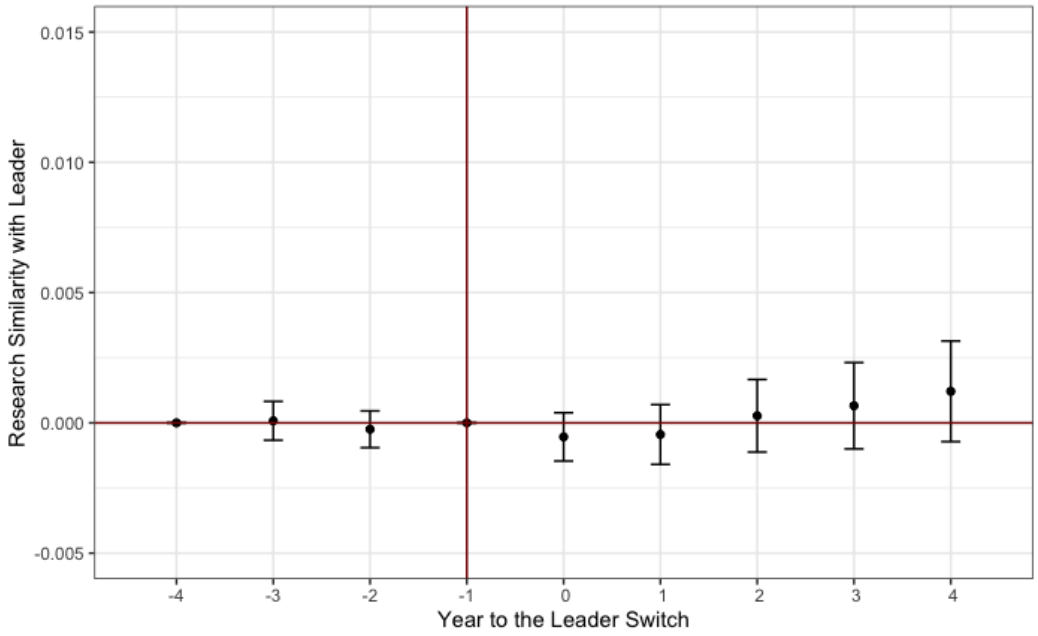


Figure 1: Impact of leader switch on the faculty-leader similarity score. The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the nonparametric event study in equation 3). The error bars represent the 95% confidence intervals.



Panel A: Impact of leaders from other higher ranked schools ($\mathcal{L}2$) vs. leaders from other lower ranked schools ($\mathcal{L}3$) on the similarity score



Panel B: Impact of leaders from other disciplines ($\mathcal{L}4$) on the similarity score

Figure 2: Impact of the switch for different types of leaders on the faculty-leader similarity score. Panel A uses the faculty-leader pairs in which leaders from other higher ranked schools ($\mathcal{L}2$) and leaders from other lower ranked schools ($\mathcal{L}3$). We estimate the effect of the two type of leaders simultaneously in regression $Y_{i,t,c} = \sum_{l\neq -1;l=-3}^{4} \mu_l^k D_{i,t,c}^l \times L_i + \sum_{t\neq -1;t=-3}^{4} \psi_l D_{i,t,c}^l + \alpha_i + \lambda_t + v_{i,t,c}$, where L_i is the indicator for whether the leader in pair i is L3 leader (=1) or not (=0). The grey lines/markers represent the estimated effects of L2 leaders (i.e., the ψ_l in the regression). The yellow lines/markers represent the estimated effects of L3 leaders (i.e., the $\mu_l + \psi_l$ in the regression). Panel B uses the faculty-leader pairs in which leaders from other disciplines ($\mathcal{L}4$). The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the non-parametric event study in equation 3).

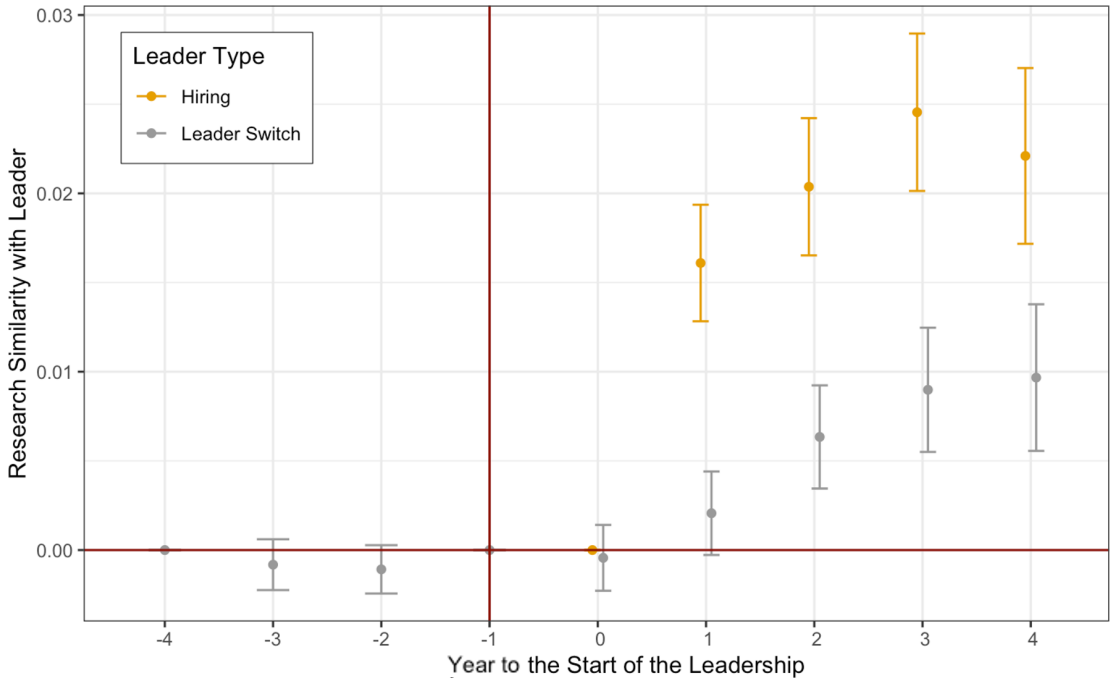
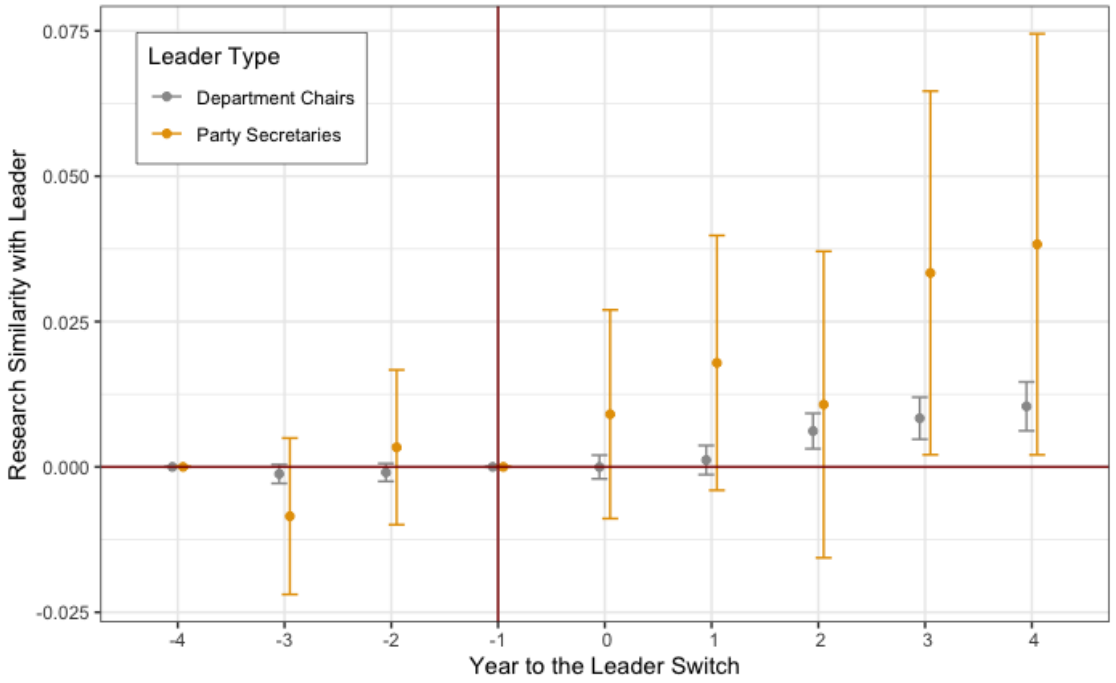
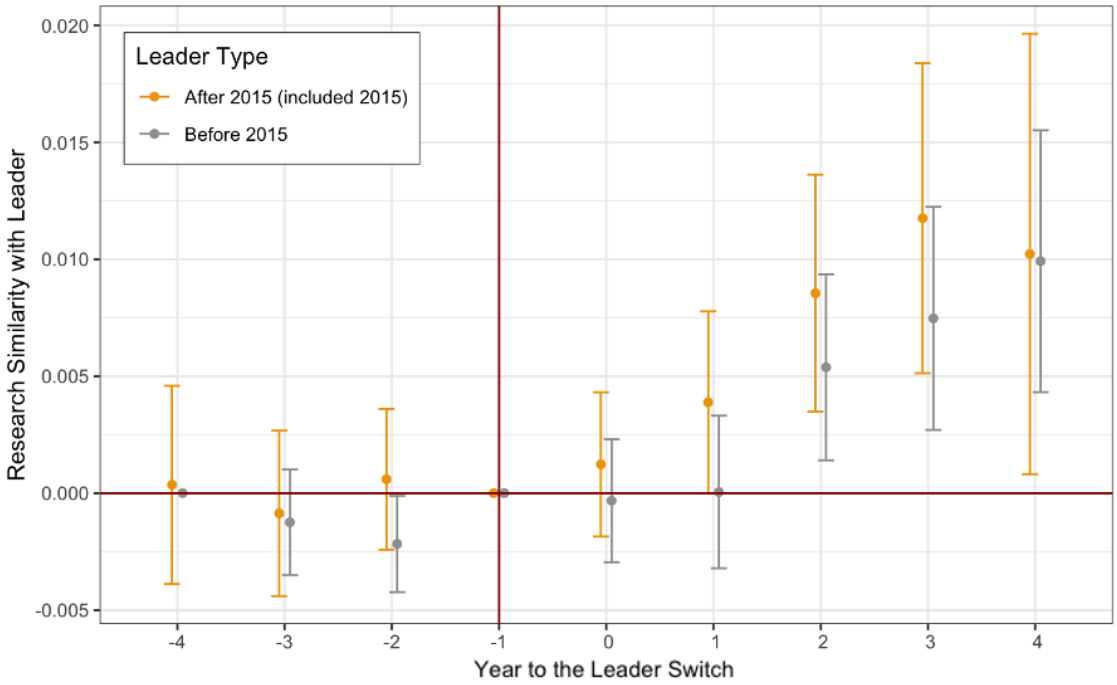


Figure 3: Impact of the leader switch vs. hiring leaders. The yellow lines/marks represent the estimators for the effect of hiring leaders. The grey lines/marks represent the effect of leader switch. We add faculty members that are newly hired to the baseline sample: (1) similarity scores will the similarity between the faculty member and the leader hired her; (2) the treatment year is the year when the faculty member is recruited by the school.



Panel A: Heterogeneous effect of department chairs and party secretaries.



Panel B: Heterogeneous effect of leader switch before and after 2015.

Figure 4: Panel A: The grey lines/marks represent the estimators for department chairs with the baseline regression. The yellow lines/marks represent the estimators for party secretaries, which is also based on the baseline regression. Since many party secretaries are non-academic personnel, we restrict our sample to party secretaries whose productivity is above the median of department chairs to ensure that we are focusing on academic party secretaries; Panel B: We estimate the effect of leader switch before and after 2015 simultaneously in regression $Y_{i,t} = \sum_{l\neq -1; l=-3}^{4} \mu_l^k D_{i,t}^l \times \text{After}_i + \sum_{t\neq -1; t=-3}^{4} \psi_l D_{i,t}^l + \alpha_i + \lambda_t + v_{i,t}$, where After_i is the indicator for whether the leader in pair i takes office after 2015 (=1) or not (=0). The grey lines/markers represent the estimated effects of leaders before 2015 (i.e., the ψ_l in the regression). The yellow lines/markers represent the estimated effects of leaders after 2015 (2015 included) (i.e., the $\mu_l + \psi_l$ in the regression).

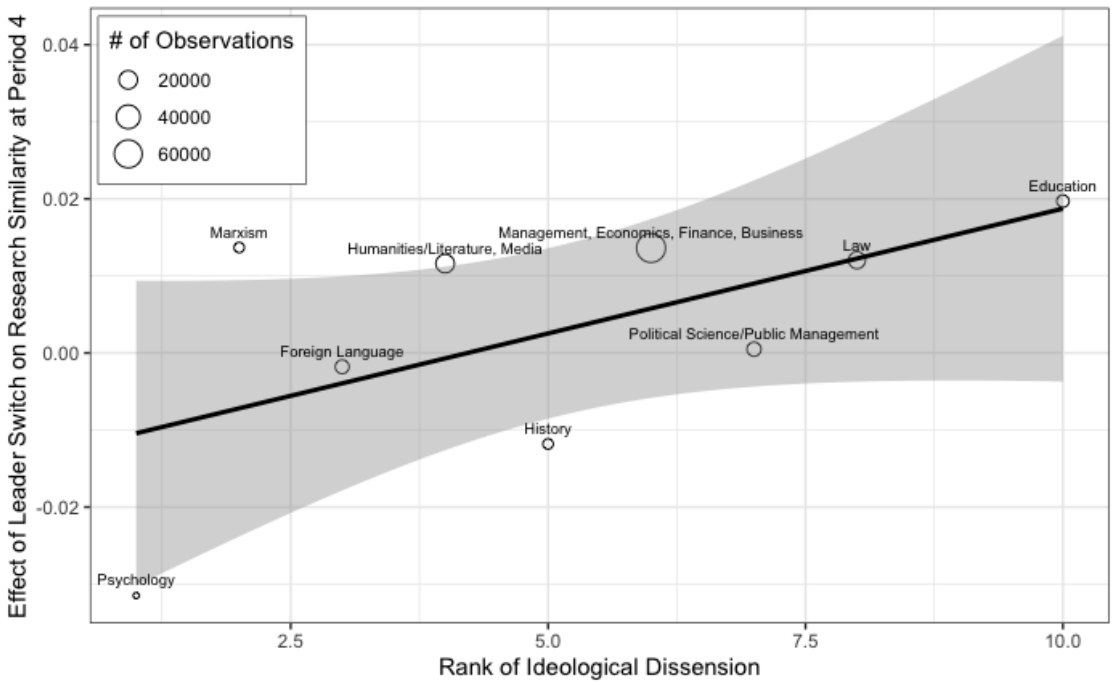


Figure 5: The correlation between the effect of leadership switch and the measure of ideological dissension. The x-axis shows the ranking of "ideological dissension." There is less ideological dissension in the discipline when the ranking is lower. The y-axis is the effect of leader switch at t=4. Each point represents a discipline.

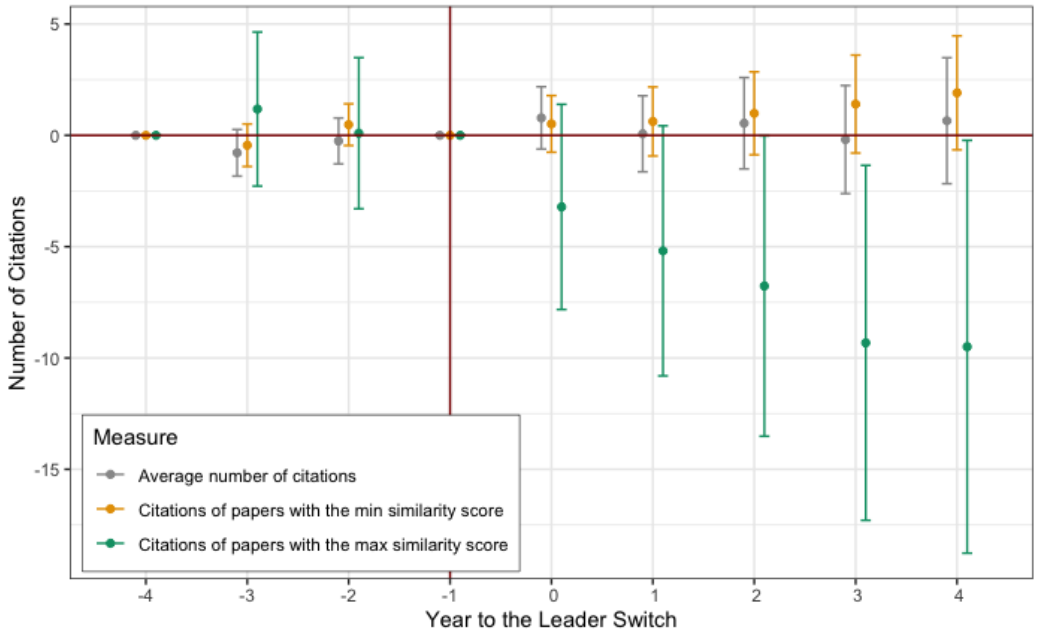


Figure 6: Impact of leaders on the number of citations. The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the nonparametric event study in Equation 3), using three measures of citations. The error bars represent the 95% confidence intervals. The grey lines/markers represent the estimated effects on the average number of citations. The yellow lines/markers represent the estimated effects of leaders on citations of papers with min similarity score. And the green lines/markers represent the estimated effects of leaders on citations of papers with max similarity score.

Table 1: Hiring leaders vs. other leaders

	Similarity Score		
	Pooled (1)	First year (2)	Pooled (3)
Panel A: TF-IDF Mean			
Dummy for hiring leader	0.00019	0.00015	0.00014
	(0.00021)	(0.00032)	(0.00021)
Panel B: TF-IDF Max			
Dummy for hiring leader	0.01097	-0.00226	0.00893
	(0.00137)	(0.00197)	(0.00139)
Faculty FE	Yes	Yes	Yes
Calendar Year FE	Yes	Yes	Yes
Control for Event Time	No	No	Yes
Number of obs	207852	61947	207852

Note: (1) Use the sample where the relative year is between 0 and 4. Regression here is: $Y_{ijt} = \beta H_{ijt} + \alpha_i + \gamma_t + \epsilon_{ijt}$, where Y_{ijt} is the similarity score between faculty i and leader j at calendar year t, H_{ijt} is an indicator for whether leader j is the hiring leader of faculty i at year t, α_i and γ_t control for the faculty member and year fixed effect.
(2) Only use the sample where the relative year equals to 0. Regression here is: $Y_{ijt} = \beta H_{ijt} + \alpha_i + \gamma_t + \epsilon_{ijt}$.
(3) Use the sample where the relative year is between 0 and 4. Regression here is: $Y_{ijt} = \beta H_{ijt} + \sum_l \psi_l D^l_{ijt} + \alpha_i + \gamma_t + \epsilon_{ijt}$, where D^l_{ijt} is an indicator for faculty i and leader j being l periods away from initial treatment at calendar year t.

APPENDIX

A Categorizing multidisciplinary schools

The categorization is done with the following steps:

Step 1 For each school, we extract disciplines that are (i) listed in school level code, and (ii) listed in the name of departments that are under the school. We need (ii) to improve accuracy, since some schools can cover disciplines that are not indicated by their school-level names. For example, the Guanghua School of Management of Peking University (北京大学光华管理学院) only has the discipline "management" in its name, but actually covers other disciplines including business, finance and economics.

At the end of this step, for each school, we get an array of disciplines in the school. For example, the Guanghua School of Management of Peking University has the following disciplines: (Management, Management, Management, Economics, Economics, Finance, Finance, Business).

Step 2 Within each school, we drop disciplines which only take less than 25%. The threshold is chosen ad hoc by checking if the final categorization makes sense. Notice that there are 10 out of 787 schools are dropped in this process since there is no discipline in the school is more than 25%. These schools are usually called "School of Social Science" (人文社会科学院) which are a mixture of all social sciences. If we include the 10 schools, all disciplines need to be combined into one.

After this step, one school has at most 3 kinds of disciplines. For example, the Guanghua School of Management of Peking University has 3 disciplines: Management, Economics, Finance.

Step 3 Within each school, disciplines are ranked by percentage. So the first discipline will be the major discipline of the school. Given the first discipline, we need to check what other disciplines are usually linked to it. Some links are very rare. For example, most of the Marxism schools are independent of other disciplines. But Northeastern University (东北大学) combines Marxism and Philosophy, and China Agricultural University (中国农业大学) combines Marxism and History. These links will mess up the categorization. So I dropped these rare links and categorize the schools with their first discipline. The criteria that we use is: given the first discipline, drop if this type of links only takes less than 15%.)

Step 4 Finally we group observations by the connected disciplines by using *group_twoway* by Mation and Maciente (2014). The final categorization is as follows:

- Marxism
- Political Science, Public Management
- Law
- Management, Economics, Finance, Business
- Education

- Foreign Language
- Humanities, Literature, Media
- History
- Psychology
- Philosophy, Anthropology, Ethnology, Sociology
- Regional Studies

B Identification of faculty members

In this section we discuss the strategy we use to identify faculty members from students or other unaffiliated researchers for a given department.

The difficulty for getting a full set of faculty from 1990 and 2019 is that most of the universities don't have good records of faculty at department level. We utilize the scientific publications of all affiliates in the 109 universities and extract faculty lists based on authors and affiliations to pin down the list of faculty and assign them to schools they are affiliated.

The general process is: *(i)* Manually find keywords for identifying department; *(ii)* Filter faculty members with certain criteria.

Manually find the "keys" for identifying department This is a key step for identifying faculty members. The affiliation entries in our publication data are typically very messy. What makes thing worse, some people don't use the full name of their departments/schools to put it in the affiliation. For example, a professor affiliated to the Department of Applied Economics of Guanghua School of Management at Peking University (北京大学光华管理学院应用经济系) could possibly put something like "Guanghua School at Peking University" (北京大学光华学院) or "DAE of Guanghua School of Management at Peking University" (北京大学光华管理学院应经系) to the affiliation of his paper. Therefore, we must manually extract some "keyword" to match the affiliation. The protocol for adding the searching keywords is:

1. University names are automatically added to the set of keywords for all the affiliated departments;
2. We make sure that each set of keywords can uniquely identify one department. One should be very careful when trying to use generic keywords to identify a department. For example, when we try to use "Fianace" (金融) to extract papers and faculty members for the Department of Finance of School of Economics at Peking University (北京大学经济学院金融系), which will be contaminated by the Department of Finance of Guanghua School of Management at Peking University (北京大学光华学院金融系) and the Department of Financial Mathematics of School of Mathematics at Peking University (北京大学数学学院金融数学系). Therefore a possible key in this case to uniquely identify the Department of Finance of School of Economics at Peking University (北京大学经济学院金融系) could be "Peking University" (北京大学), "Econ" (经), and "Finance" (金融).
3. Most affiliations in the papers are precise only to the school level (just like HBS and Harvard SEAS), not to department level (in US equivalent, Harvard Econ department and Political Science department). This is because, most Chinese "schools" are US "departments" equivalent, and Chinese "departments" are equivalent to something like the macroeconomics group at Harvard econ department, although there are typically still bureaucracy structures and CCP establishments in this very disaggregated level. We try to identify faculty members by their department (in the Chinese sense), however in many times we are not able to do that.

Filter faculty members with certain criteria After we can identify department/school from papers, we use these to identify faculty members. Our current criteria for a faculty member are: *(i)* having more than 3 years of publication span; *(ii)* having 5 or more papers. The first is to exclude PhD students who typically publish papers in 2-3 years, and the second is to guarantee enough variations for us to exploit in regressions.

Validation with one university To validate the method that we use to extract faculty members, we compare the faculty we extracted with the faculty list that we can obtain from the the official website of the School of Economics at Sun Yat-sen University. Table A.1 shows the validation result.

Table A.1: Compassion between extracted faculty and the actual personnel

Year	Actual Number of Faculty	Number of people extract from raw data	Number of Faculty after filtering
2019	44	353	23
2018	41	327	31
2017	37	305	35
2016	35	301	39
2015	33	297	39
2014	31	296	40
2013	29	337	44
2012	28	335	46
2011	24	340	52

C Similarity score based on Doc2Vec methods

Possible measurement error in TF-IDF The biggest issue for TF-IDF is that it can severely underestimate the similarity between two documents, which will be a major measurement error for our baseline regression. Consider comparing the similarity between the following two documents:

- Document A is "The car is driven on the road."
- Document B is "The truck is driven on the highway."

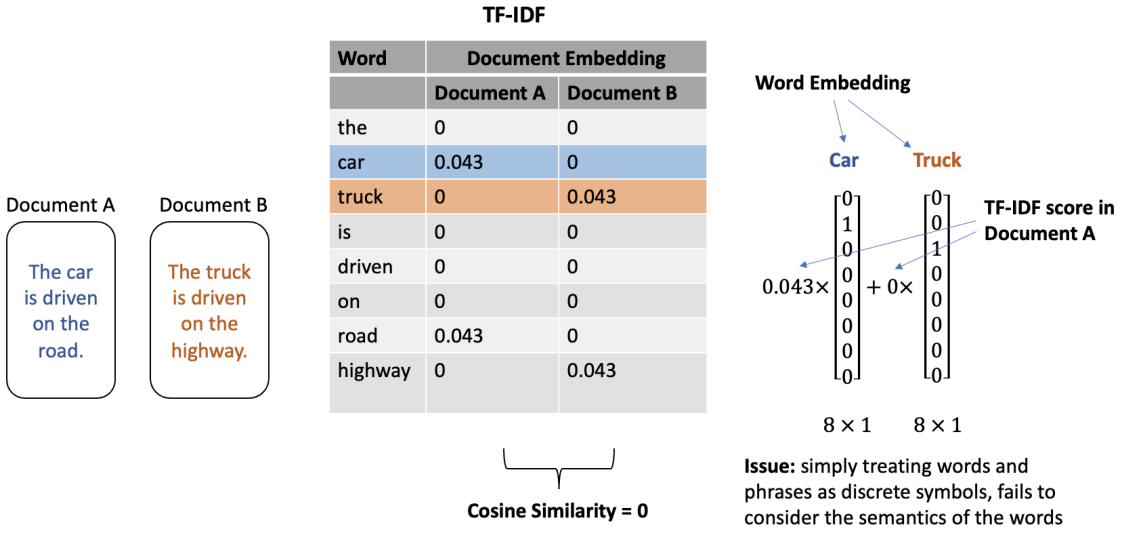


Figure A.1: A illustrative example of TF-IDF

The table in the middle shows the document embeddings (vectors) of the two documents. The document embedding is an 8x1 vector, where 8 is the size of the vocabulary. Each element in the vector is the TF-IDF score of the corresponding vector. After calculating the cosine similarity score of the two documents, we get the similarity score between the two documents equals 0. For each document embedding, we can view it as a weighted average over all of the word embeddings in the document. As shown on the right of Figure A.1, the weights are TF-IDF scores, and the word embeddings are standard unit vectors.

Representing words as standard unit vectors treats words as discrete symbols. We assume "car" and "truck", "road" and "highway" are totally different, which leads to the result that the similarity score is 0.

Taking semantics of the words into consideration Replacing word embeddings with Word2Vec (Mikolov et al., 2013) will relax the assumption of TF-IDF that words are orthogonal.

The Word2Vec algorithm uses a neural network model to learn word associations from a large corpus of text. One of the Word2Vec structures – continuous bag of words (CBOW) – forms a sliding window around the current word and allows it to be predicted using the

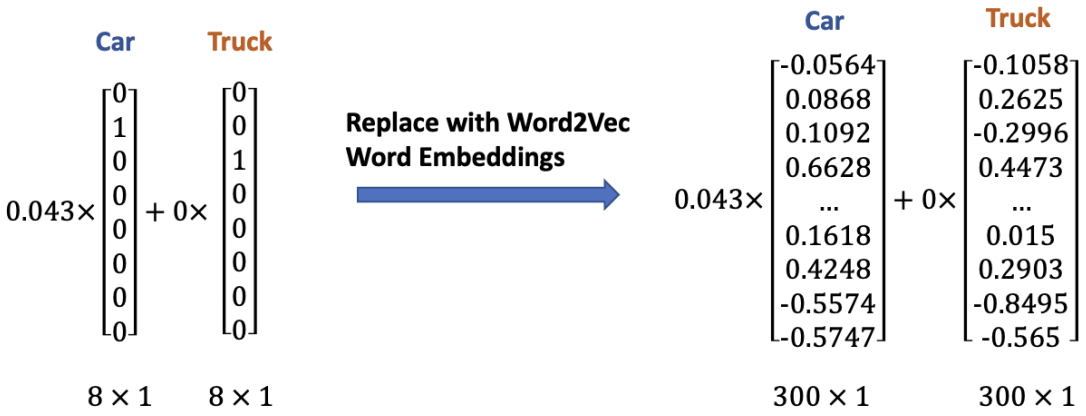


Figure A.2: Change from TF-IDF to Weighted Word2Vec

words around it. A feature vector is used to represent each word. With this process, each word will be mapped to a $n \times 1$ vector and can represent the semantic similarity between words[1]. If we reconsider our previous example, using word embeddings from Word2Vec will increase the similarity score between Document A and B. The cosine similarity between "car" and "truck" is 0.81 rather than 0 in the TF-IDF case. With the new weighted average over all of the word embeddings in the document, the similarity score between Document A and B is 0.8442.

While training an informative Word2Vec model requires large train set and computation power, many researchers and institutions publicize their training results based on large corpus such as Wikipedia, newspapers, and twitter. One can use their results as pre-trained word embeddings and further finetune models to fit her own target corpus.

In our setting, we utilized two pre-trained Chinese word embeddings provided by Li et al. (2018): *(i)* Word2Vec trained with People's Daily News, and *(ii)* Word2Vec trained with Chinese Wikipedia.

Taking the ordering of the words into consideration Although weighted Word2Vec relaxes the orthogonality of semantics, another assumption is still concerning when numerically representing paragraphs – it ignores the syntax within each paragraph. Doc2Vec is developed to further tackle this issue and give paragraph embeddings.

Instead of directly using TF-IDF weights, Doc2Vec puts more structure on the relationship between words and documents to estimate the weights. Based on CBOW which utilizes words to anticipate the following word, Doc2Vec also includes a document-unique feature vector. As a result, when the word vectors are trained, the document vector is also trained which can be used as the numeric representation of the document.

[1] n will be a hyperparameter in the Word2Vec model. In the example here, we set $n = 300$.

D Additional figures and tables

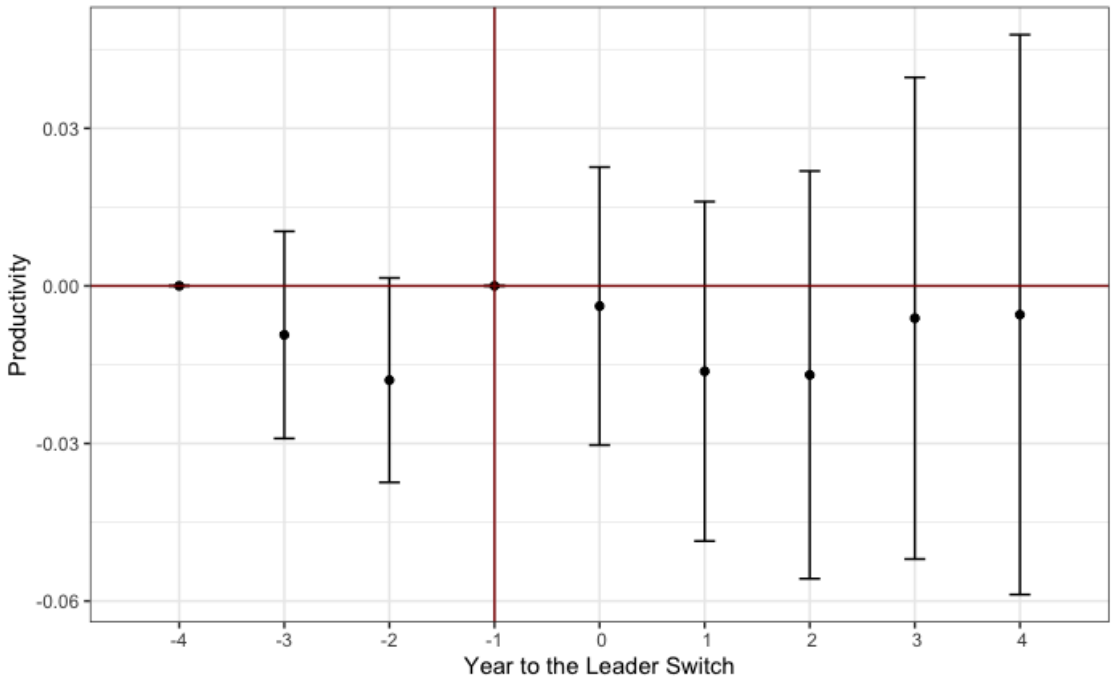


Figure A.3: Impact of leader switch on the productivity of faculty. Using the number of publications per year as the outcome variable, we re-estimate the baseline specification (equation 3) restricting the sample to a balanced panel of faculty-leader pairs.

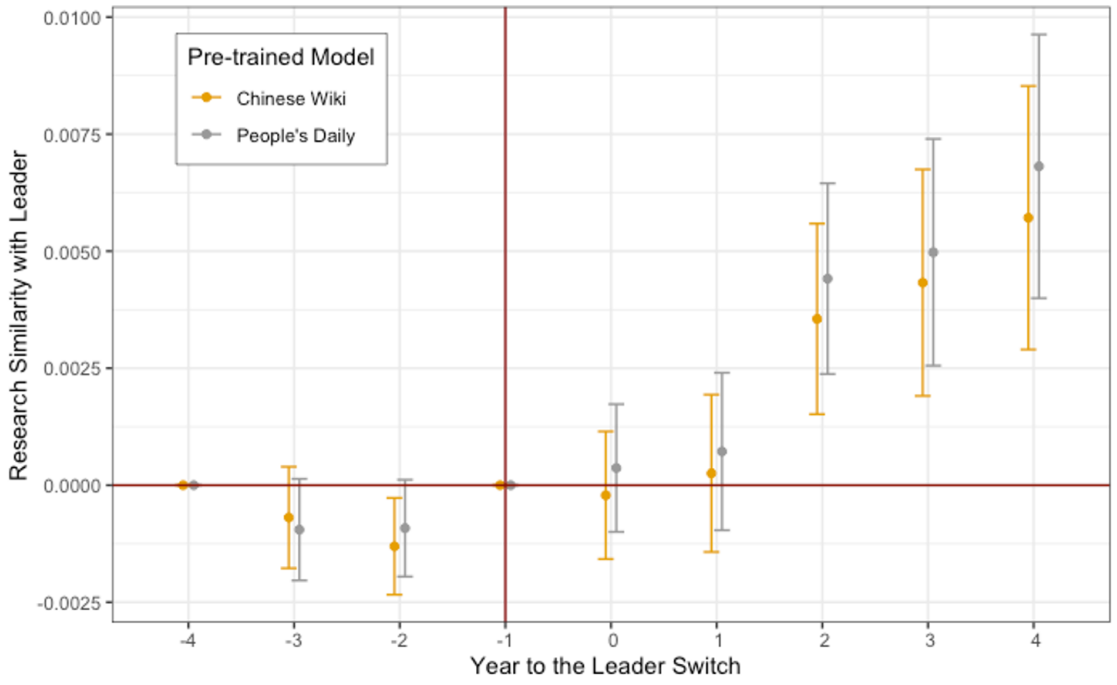


Figure A.4: Robustness check: impact of leader switch on the faculty-leader Doc2Vec similarity score. The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the nonparametric event study in Equation 3). The error bars represent the 95% confidence intervals.

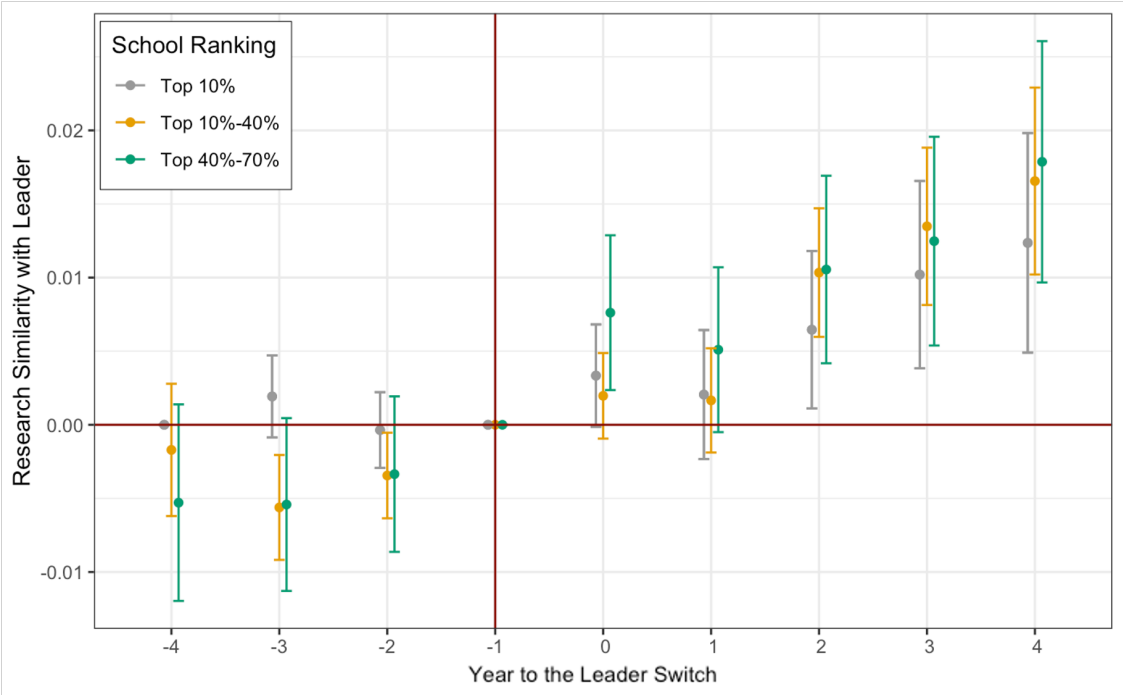


Figure A.5: Heterogeneous effect of leader switch by school ranking. We estimate the effects simultaneously in regression $Y_{i,t} = \sum_k \sum_{l \neq -1; l=-3}^{4} \mu_l^k D_{i,t}^l \times R_i^k + \sum_{t \neq -1; t=-3}^{4} \psi_l D_{i,t}^l + \alpha_i + \lambda_t + v_{i,t}$, where R_i^k is the indicator for the rank of the school of pair i. The grey lines/markers represent the estimated effects of leaders from schools ranked top 10%. The yellow lines/markers represent the estimated effects of leaders from schools ranked 10%-40%. And the green lines/markers represent the estimated effects of leaders from schools ranked 40%-70%.

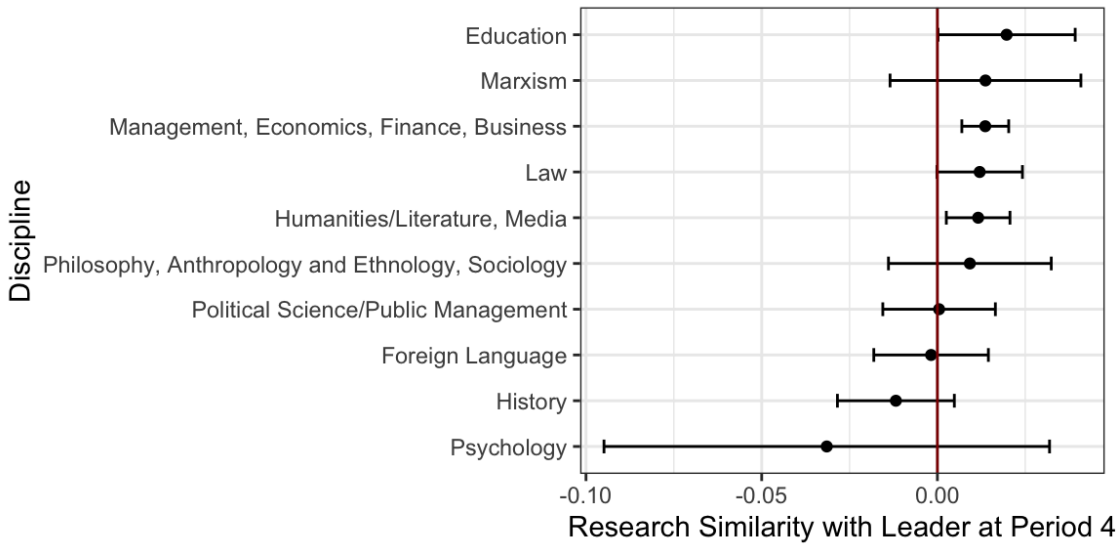
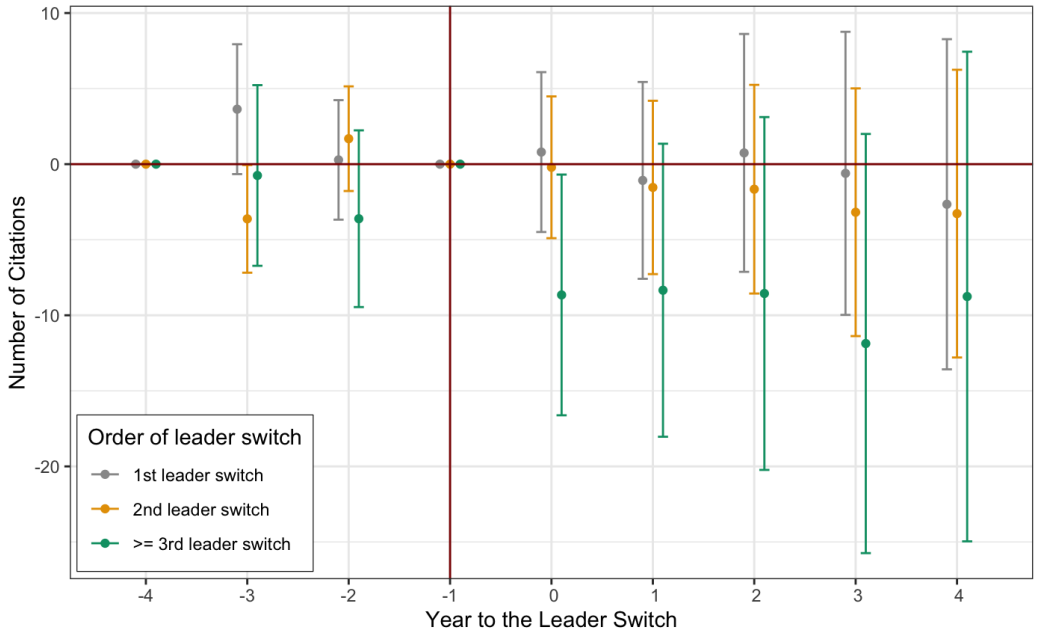
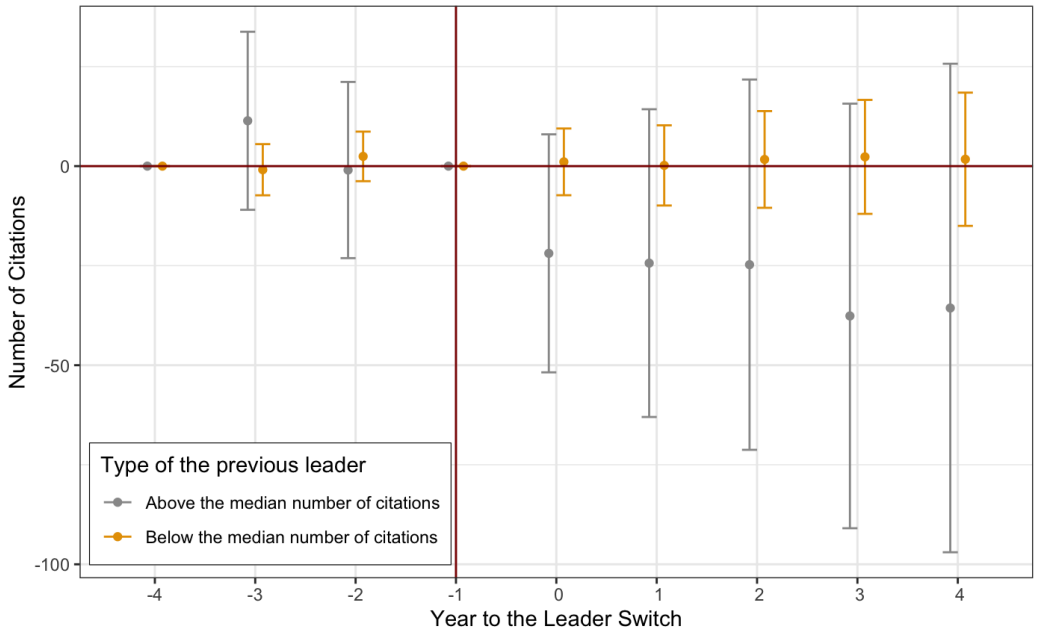


Figure A.6: Heterogeneous effect of leader switch by discipline. We estimate equation 3 for each discipline. Each line/mark represent the point estimate of the 5th year for each discipline, ranked by size of the estimates. We classify schools into 10 categories by taking disciplines as the connected components of schools that share the same classification. The details are described in Appendix A.



Panel A: Heterogeneous effect by the order of leader switch.



Panel B: Heterogeneous effect by the type of previous leaders, given current leader is below the median number of citations.

Figure A.7: The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the nonparametric event study in Equation 3), using three measures of citations. The error bars represent the 95% confidence intervals. Panel A: The grey lines/markers represent the estimated effects for leader switches that are firstly experienced by faculty. The yellow lines/markers represent the estimated effects of the second leader switch experienced by faculty. And the green lines/markers combine the effect of the rest of the switches; Panel B: The grey (yellow) lines/markers represent the estimated effects for leader switches from above (below) median quality leaders to below median quality leader. The median quality is defined as the median of the average citations before leaders take office. The quality of the citation data might be different for different years. For early years, since all papers are in print, the citation record might not be complete for them. This will make leaders in the 1990s always below the median if we don't control the secular variation. So we calculate the median for each discipline and decades instead.

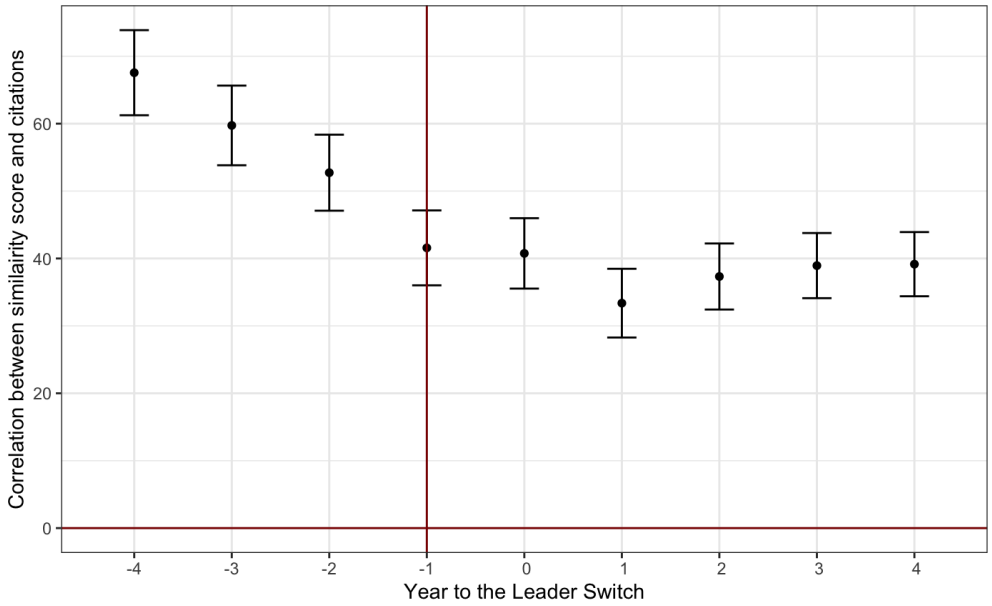


Figure A.8: Correlation between Similarity score and the number of citations. The points in the figure represent the coefficients ψ_l in the equation
Citation $_{i,t} = \sum_{l \neq -1; l=-4}^{4} \psi_l D_{i,t}^{l} \times$ Similarity $_{i,t} + \alpha_i + v_{i,t}$, where Citation $_{i,t}$ is the average number of citations for faculty i at year t, Similarity $_{i,t}$ is the TF-IDF Max similarity score between faculty i and her leader at year t, α_i is faculty FEs.

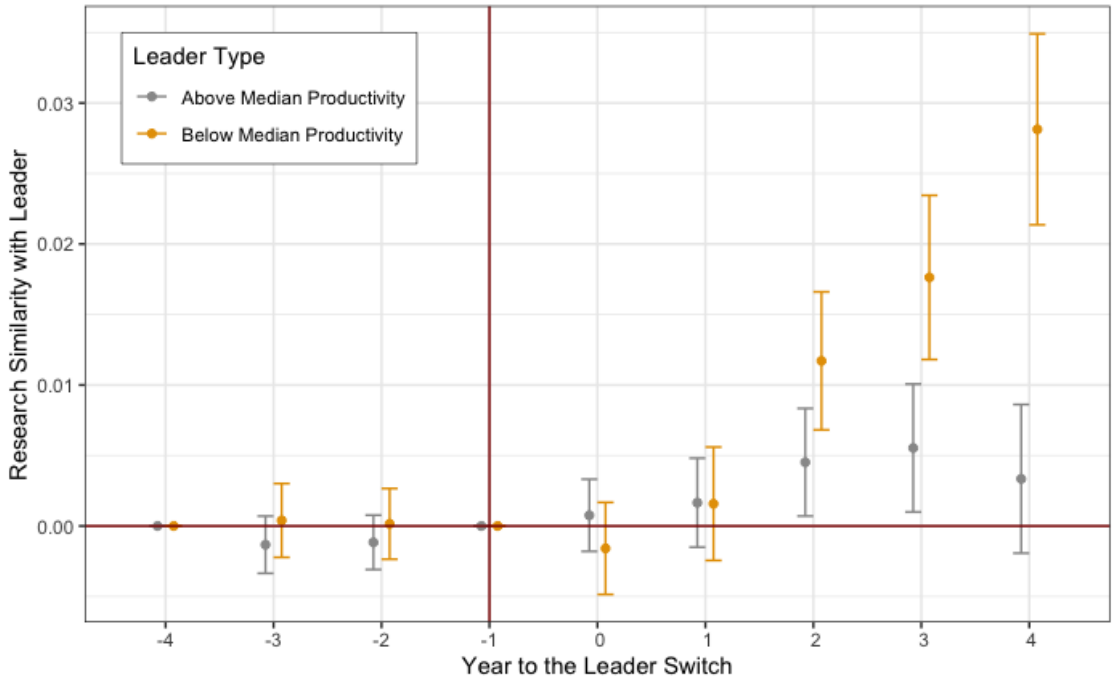


Figure A.9: Heterogeneous effect of leader switch by the productivity of leaders. We define the productivity of a leader as the average number of publications of the leader in the 3 years before she is in office. Then we take the median of leaders' productivity for each discipline. A leader will be categorized as "above mean productivity" if her productivity is above the median number of her discipline. Otherwise, she will be categorized as "below median productivity". We estimate the effect of leader with different productivity level separately with our baseline regression (i.e., we assume that pairs with different leader productivity have different calendar year fixed effects). The yellow lines/marks represent the estimators for the effect of leaders that are below the median productivity. The grey lines/marks represent the estimators for those above the median productivity.

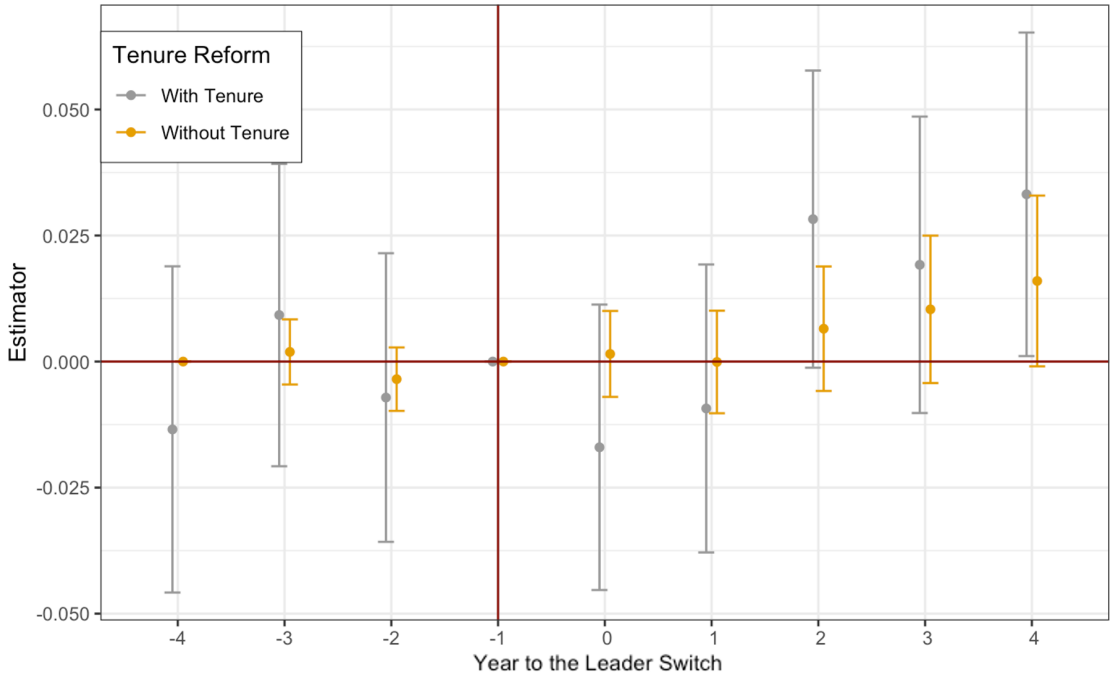


Figure A.10: Heterogeneous effect before and after tenure track reform at Peking University. The points in the figure represent the estimated effects of event time (i.e., the ψ_l from the nonparametric event study in Equation 3), using three measures of citations. The error bars represent the 95% confidence intervals. The grey (yellow) lines/markers represent the estimated effects for leader switches for faculty-leader pairs that are affected (unaffected) by the tenure track system reform.

Table A.2: Summary Statistics on Paper Citations

	Citations per Paper		
	All Papers	Faculty Only	Leaders Only
Mean	13.9	13.7	18.2
25th Percentile	1.0	1.0	1.0
50th Percentile	4.0	4.0	5.0
75th Percentile	13.0	13.0	16.0
Total Papers	736,756	725,455	115,235