THE UNIVERSITY OF CHICAGO

THREE ESSAYS ON ELITE AND POPULAR POLITICS AND CULTURE

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF SOCIOLOGY

BY
SHILIN JIA

CHICAGO, ILLINOIS
AUGUST 2022

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I have been in debt to too many people in my life.

The first person I must thank is my dissertation committee chair and Ph.D. advisor Dr. John Levi Martin. John is unarguably the best mentor in the world. He has always been willing to invest tremendous amount of time and energy into mentoring. And if there is any person in the world who truly understands who I am as an intellectual person, this person is going to be John. He understands the inner drive of my projects from the very beginning and has always been supportive despite my numerous failures. I am not a person who is always good at expressing my ideas to my peers. But John oftentimes knows exactly what I am thinking even when I am still struggling in putting my words together. And it is usually the case that he can articulate my ideas much better than I can. It is such a great fortune to have him as my mentor in my career.

This dissertation would also have not been possible without the generous support of my committee members. Dr. James Evans has basically been my de facto co-advisor since I took his classes in year 1. He is the person who introduced me to computational social science and has taken me under his wings since then. It was such a defining moment of my career. And even though James has always been super busy all the time, he has consistently been the person who I can rely on during moments of crisis. He has saved my career many times.

I am also in great debt to my other two committee members, Dr. John Brehm and Dr. Junyan Jiang. As a sociologist who studies politics, I have the fortune to have two great political scientists in my committee. John and Junyan both generously agreed to serve on my committee not because of our limited previous interactions but because of our shared academic interest. I am especially thankful to them for their willingness to tolerate my mistakes.

I would also have to express my deepest gratitude to my collaborators, Dr. Benjamin Rhor and Dr. Linzhuo Li. Chapter 1 and Chapter 2 of my dissertation are the results of our

collective intellectual journey. During the collaboration, Benjamin demonstrates his magic by saving my vacancy chain project from my hard disk. I have been constantly impressed by his remarkable analytical skills in questioning and strengthening my vacancy chain arguments. And Linzhuo is always the source of creativity in my People's Daily project. The bold journey Linzhuo and I took was only the result of our countless days and nights of debates and discussion.

Then, I have to mention the generous support I received from the sociology department and especially Dr. Kristen Schilt, the Director of Graduate Studies of my home department. I feel privileged to have been able to spend 8 years of my life engaging in intellectual activities that I otherwise wouldn't be able to do in any other university in the world. I am especially grateful for the Henderson Research Grant I received in the last year of my study. The grant allowed me to hire three brilliant undergraduate research assistants, Meng Zheng, Kejia Li, and Hangtao Li, to complete the data collection project that I had previously spent years working on alone. I would like to thank all the administrative staff members in my department and especially Linnea Martin. Their dedication has allowed me to focus on my academic work and oftentimes forget about the existence of bureaucracy.

I am also thankful to my secondary affiliation, the Master's Program in Computational Social Science (MACSS). The program offers me a home and a secured working environment during my advanced years. The preceptor job I took also helped me accumulate tremendous amount of learning experience in transitioning from a researcher to a teacher. I am especially thankful to Dr. Chad Cyrenne, Dr. Benjamin Soltoff, and Dr. Sanja Miklin for their help on my teaching career.

Outside the University, I am grateful to have two great sociology mentors in my life, Dr. Ray-may Hsung and Dr. Horng-Luen Wang. Although I don't have much experience interacting with them in recent years, the time when they introduced me to sociology always remains the most valuable memory in my life and continuously motivates my ongoing

research.

Last but not the least, the greatest debt I owe is always to my family. As always, my parents, Guorong Jia and Jianyuan Chen, lend me unlimited support without asking for any return. I am also lucky enough to have met and married my wife, Meiwen Lu, during my Ph.D. years. The moments of joy I have with her makes my life never the same again.

ABSTRACT

This Ph.D. dissertation consists of three independent essays on elite and popular politics and culture. Chapter 1 is a study of the career mobility of high-level political elites in post-reform China. The essay develops an account that views mobility not from the perspective of the individual, but from the perspective of the organization trying to fill positions. Our study shows that during the reform era elites whose transfers are embedded in vacancy chains are more successful than those whose transfers occur in isolation. These findings are the result of a strategy of organizational sponsorship pursued by the Communist Party of China that results from efforts to integrate the increasingly decentralized Chinese state. Chapter 2 is an essay that focuses on the ideological dimension of the Chinese reform. We found that during the course of China's economic reform, the Party's official state rhetoric progressed path-dependently at a highly consistent pace. Without being absorbed internally, external political shocks could quickly die out in the system. New elements were likely to survive only in existing stable contexts. The results suggest that in cultural production, adaptation in stable configurations is a key step in long-term cumulative change. Chapter 3 is a study of how books in 6 major languages of the world perceive other nations and identities when their host nations are in war and trade with other nations. We show that international perceptions change significantly during war time. And the patterns are consistent in American English, British English, and French.

CHAPTER 1

VACANCY CHAINS AS STRATEGY: INTER-ORGANIZATION MOBILITY OF POLITICAL ELITES IN REFORM CHINA

Shilin Jia and Benjamin Rhor

Chapter Abstract Questions of career advancement and success have long occupied scholars of the Chinese state. Most work on political mobility in China views mobility from the perspective of the individual. It imagines a pool of officials who are at risk of moving up a status hierarchy of positions and the goal is to explain why some succeed and others fail. We develop a complementary account that views mobility not from the perspective of the individual, but from the perspective of the organization trying to fill positions. Such a perspective does more justice to the bureaucratic nature of the Chinese party-state in which hiring is centrally coordinated by high-level elites in the CCP. Focusing on the organization directs our attention to the strategies behind personnel management. We argue that those strategies can be studied through an analysis of vacancy chains. Drawing on a novel dataset of more than 2,500 inter-organizational transfers derived from the CVs of more than 4,000 political elites included in the Chinese Political Elite Database, we show that between 1977 and 2012 elites whose transfers are embedded in long vacancy chains are more successful than those whose transfers occur in isolation. In addition, we demonstrate that this career boost occurs after their involvement in vacancy chains and that it is stronger for younger elites than for older ones. We contend that these findings are the result of a strategy of organizational sponsorship pursued by the CCP that results from efforts to integrate the increasingly decentralized Chinese state.

INTRODUCTION

Dynamics of job mobility and status attainment have long been major objects of sociological inquiry (Blau and Duncan 1967; Lin 1999; Rosenfeld 1992; Kalleberg and Mouw 2018). Most existing work that seeks to explain career advancement and career success views mobility from the perspective of the individual. It imagines a pool of individuals who are at risk of moving up a formal or informal status hierarchy of positions and the goal is to identify individual-level factors that explain why some individuals eventually reach high positions while others do not. This is true both for studies that understand mobility as a function of individual characteristics such as human capital (e.g. Coleman 1988; Lin 1999) and for studies that incorporate the larger opportunity structure of available positions into individual-level models (e.g. Sorensen 1974, 1977). This individual-centered view of mobility is related to the notion of a career as a sequence of positions through which an individual moves (Spilerman 1977).

But this is only half the story. Job mobility always has two sides: the person striving to be chosen and the person or organization making the choice. Hence, a dual, less common perspective starts not with the individuals who occupy the positions, but with the organizations responsible for filling them. Here mobility appears as the result of vacancies in the system and the process by which these vacancies are filled. This organization-centered view of mobility is related to the notion of staffing as a process of filling vacant positions. While this perspective is not commonly adopted in studies of mobility, it is familiar to organizational scholars (H. Kaufman 1960; Kanter 1993; Pinfield 1995; Taylor and Collins 2000) and scholars of political elites (Moe 1985; Gallagher and Marsh 1988; Lewis 2008, 2011; Van Gunten 2019). Adopting the organization-centered perspective directs attention to the strategies that the organization pursues through its staffing practices.

These two perspectives are related to two distinct processes of allocating individuals to positions and positions to individuals (Turner 1960; Padgett 1990; Putnam 1976). The first

process assigns individuals to positions based on competition in a market. Here high status is the result of an open contest in which individuals compete using the skills and credentials they have acquired. The second process assigns individuals to positions based on centralized management. Here high status is the result of the strategies pursued by a central planning authority.[1]

In some cases, it is reasonable to view mobility in terms of individual merit and striving. In studies of mobility at the level of national or even international labor markets, where staffing decisions are distributed over many different organizations and career moves are largely independent, the person-centered perspective is often appropriate. In other cases, however, it may lead us to miss the most important dynamics governing career advancement. This is likely to be the case in analyses of mobility within organizations where the authority over hiring decisions is centralized in an organizational subunit or a small number of individuals. Here we can gain new insights by focusing on what guides the decisions to advance people and the organizational strategies behind hiring and promotion.

Due to its bureaucratic nature, the Chinese party-state is one such case (Huang 2002). Because the number of positions in the Chinese bureaucracy is relatively fixed, a person's ability to move is constrained by the availability of vacancies in the system. In addition, the process of filling high-level vacancies is centrally organized and controlled by elites within the Chinese Communist Party (CCP) who regard control of personnel as an important pillar of one-party rule (McGregor 2010; Edin 2003; Landry 2008; Zheng 2010). This creates a system in which careers are a function not only of the ambitions and preferences of officials but of the demands of the state and the strategies pursued by the CCP.

1. Turner uses this distinction in a comparative study of social mobility in the American and English school systems. The American system is characterized by what he calls "contest mobility." In such a system, "elite status is the prize in an open contest and is taken by the aspirant's own efforts." The British system, in contrast, is one of "sponsored mobility" in which "elite recruits are chosen by the established elite or their agents, and elite status is given on the basis of some criterion of supposed merit and cannot be taken by any amount of effort or strategy. Upward mobility is like entry into a private club where each candidate must be 'sponsored' by one or more of the members" (Turner 1960, 856). Putnam (1976) and Li and Walder (2001) apply this distinction to political elite recruitment.

Yet, similar to work on status attainment in general, much of the work on political mobility in China also adopts the person-centered view. It seeks to understand the career advancement and ultimate career success of political officials from the perspective of the individuals striving for high office. Early work in this tradition has looked at the relative importance of party loyalty on the one hand and meritocratic elements such as educational credentials on the other (Walder 1985; Walder, Li, and Treiman 2000). More recent work focuses on officials' job performance (Maskin, Qian, and Xu 2000; Bo 2002; Li and Zhou 2005; Chen, Li, and Zhou 2005; Landry, Lü, and Duan 2017) and social capital in the form of relationships to elites high up in the party-state (Shih, Adolph, and Liu 2012; Zhang 2014; Opper, Nee, and Brehm 2015; Jiang 2018; Lu 2018; Fisman et al. 2020) that tends to get lost in this work is the side of the organization that hires and manages those officials.

In contrast to existing work, we develop an account that begins with the organizational strategies pursued by the CCP. Our goal is not to test which perspective has more explanatory power but to suggest that we can produce new insights by supplementing the individual-centered perspective with an organization-centered one that studies the organizational strategies behind hiring decisions. The central task then is to identify what those strategies are.

Because hiring decisions concerning high-level officials in the Chinese state take place behind closed doors, the strategies behind them cannot be directly observed. Work by scholars taking an institutionalist approach has therefore tried to infer them from patterns in officials' transfers (C. Li 2004; Bo 2004, 2010). Following this line of work and adopting Chandler (1962, 14)'s dictum that "structure follows strategy," we assume that the CCP's strategies leave traces in the structure of the mobility system that we can observe and use to infer the strategies that produced them. In other words, we use the patterns produced by the moves of officials through the system of positions to inform our understanding of the organizational strategies that govern them (analogous to using a medical contrast material

that lights up the structure of the body), which, in turn, will help us better understand individual career success. The person-centered and the organization-centered perspectives inform each other.

Following Pinfield (1995) and Yeung (2007), we use an analysis of vacancy chains (White 1970; Chase 1991) as a way to get at organizational strategies. The idea of a vacancy chain is counterintuitive but simple. A transfer out of the system of positions (e.g., due to retirement) generates an initial vacancy in position A, which is then filled by someone from position B, creating a new vacancy in position B, and so forth until the vacancy leaves the system of positions under investigation. Vacancy chains represent the movement of vacancies through that system of positions. If we understand vacancy chains as bundles of hiring decisions, we can treat them as residues of the personnel management strategies that gave rise to those decisions. Focusing on transfers between organizational units in the Chinese party-state, and drawing on the work of organizational scholars and scholars of the state, we identify three potential strategies—promotion based on diverse experience, on sponsorship, or on control—and show that each generates different expectations for what we should observe in the data. In particular, we are interested in whether being involved in longer vacancy chains that move across organizational units is related to career success. A strategy that values only diverse job experience suggests a positive relationship between the number of transfers an official makes and his or her career success. But beyond that, it does not matter whether those transfers are part of longer vacancy chains. Organizational sponsorship focuses on identifying promising officials early on in their careers and incorporating them into the political system by placing them on career tracks that will ultimately lead to leadership positions. Under that strategy, we would expect a positive relationship between being involved in vacancy chains and career success, and we would expect that younger people benefit more from being involved in vacancy chains than older ones. Organizational control is about preventing the formation of local pockets of power within the organization

that develops interests distinct from those of the organization as a whole. If vacancy chains are the result of a strategy of rotating officials between subunits of the organization in order to prevent them from becoming too powerful, we would expect a negative or no relationship between being involved in vacancy chains and career success, and we would expect that to the extent officials get involved in vacancy chains, they do so at the height of their careers.

To identify which of these strategies is being pursued, we draw on a novel dataset of inter-organizational transfers of Chinese officials covering the period from 1977 to 2012. We apply computer-assisted coding to the CVs of the 8172 high-level Chinese political elites included in the Chinese Political Elite Database, the most comprehensive data source of high-level CCP elites that exists. For each elite, we code all organizational units he or she worked in, which allows us to study the inter-organizational movement of both people and vacancies.

To anticipate our results, in line with previous research (Kou 2010; Bo 2010; Jia and Xu 2018), we find evidence that diverse experience, as measured by the total number of transfers an official has made, is related to career success. But on top of that, we find that both the number of times an elite has been involved in vacancy chains as well as the average length of those chains are positively related to career status. In addition, we show that the career boost associated with being part of vacancy chains occurs after an official's involvement in a chain and is higher for younger elites than it is for older ones. Being involved in a long vacancy chain indicates that an official has been identified as promising and placed on a career track that will ultimately lead to the highest positions, thus providing evidence of a sponsorship strategy.

The paper demonstrates the usefulness of a structural perspective of mobility that begins not with the individuals who fill positions but with the institutions that have been put in place for filling them. In the case of the Chinese party-state, mobility is the residue of an orchestrated strategy of sponsorship that can be seen as a governmental technique to manage

the centrifugal and centripetal forces within the vast organizational structure of the state. Individual careers, then, appear as the product of decisions that are made by high-level elites within the CCP. Changing the perspective in this way reveals the extent to which mobility systems are determined by factors that lie beyond the individuals. It is this view that takes the nature of bureaucracy and the elites who staff it seriously.

POLITICAL MOBILITY OF CHINESE ELITES

How officials move through the set of political offices within the party-state and who succeeds and who fails are questions that have long captivated observers of the Chinese state. After studying mobility as a pillar of communist rule, scholars later became interested in how the reforms that were launched in 1978 affected that system. Since 1978 the CCP has gradually transformed itself from a revolutionary party to a bureaucratic state machine that increasingly relies on a pool of career-oriented professional politicians instead of ideological zeal (Lee 1991). In addition, China has become one of the few authoritarian states in the world that have successfully implemented an institution for leadership transition and the peaceful transfer of power (Nathan 2003).

Much of the extant work on political mobility in China adopts a person-centered view. It seeks to understand career advancement and career success by relating them to individual-level characteristics. At the core is an attempt to study the relative importance of party loyalty on the one hand and meritocratic elements on the other—a distinction that has been central to work on elite recruitment more generally (Stewart et al. 1972; Putnam 1976; Lewis 2008). Early work by Walder and colleagues, for instance, looked at party membership and educational credentials and found a structure of dual career paths. Careers that lead to professional positions are based on educational credentials, while careers leading to elite administrative positions are based on party loyalty (Walder 1995a; Walder, Li, and Treiman 2000). More recent work has been particularly interested in job performance on the merit

side (Maskin, Qian, and Xu 2000; Bo 2002; Li and Zhou 2005; Chen, Li, and Zhou 2005; Landry, Lü, and Duan 2017) and social capital in the form of ties to party elites on the party-patronage side (Zhang 2014; Opper, Nee, and Brehm 2015; Jiang 2018; Lu 2018; Fisman et al. 2020). In addition, it has been shown that homophily between high- and middle-level elites plays an important role in the recruitment process of provincial leaders (Opper, Nee, and Brehm 2015). Shih, Adolph, and Liu (2012) provide one of the most comprehensive studies of political mobility among Chinese elites, examining the relative importance of most of these factors. They find that factional ties to party leaders and educational credentials contribute to career success, while other factors such as the ability to collect revenue at the provincial level, the ability to generate economic growth, time in the CCP, gender, ethnicity, and being a child of a senior leader have only modest effects or none at all.[2]

Mobility always has two sides, however: the person striving to be chosen and the person or organization making the choice. The organizational side gets lost in many of these accounts. Focusing merely on the characteristics and motivations of individuals is like trying to understand a chess game by studying the trajectories of individual chess figures. Doing that, the analyst learns a fair amount. For example, she understands that the knight has the ability to move two squares vertically and one square horizontally or two squares horizontally and one square vertically. But how is it that the knight is in a particular position at a particular point in time remains underdetermined. From the perspective of a single piece, the moves will appear random. Accounts that incorporate social capital go beyond the study of individual figures and try to understand their strength in relation to other figures in the field. They thus project structure into the individual. Our approach, which complements the

2. Even party loyalty and patron-client relationships are often viewed from the perspective of individuals who use them as a competitive advantage in a contest for positions, and not from the perspective of the state. Walder (1995a) points to two potential mechanisms that could produce a positive relationship between party membership and career success: a screening process in which party membership signals loyalty and a process of incorporation in which people are drawn into the party early on and are then groomed within it. Thus, party membership can be seen as both a credential in a contest and as a means of sponsorship. We recognize the former but our analysis emphasizes the latter.

existing ones, brings in the strategy pursued in the game. Similar to a chess commentator, we argue that an analysis of the sequential moves of the chess figures can give us insight into the strategies pursued by the players.

The organization-centered perspective promises new insights into the functioning of the Chinese political system for three main reasons. First, because the number of positions in the Chinese bureaucracy is relatively fixed, the moves officials can make between positions are constrained by the availability of vacancies. This fact is not systematically captured by analyses that focus only on individuals. Second, the authority to appoint people to positions is centralized in the CCP. This means transfers are less a function of the preferences and ambitions of officials and instead are determined by system requirements. Finally, the dramatic changes initiated by the 1978 reforms were accompanied by changes in the personnel management strategies the CCP pursued. It may be useful, therefore, to take these strategies into account to more fully understand the structure of political mobility in China.

Institutionalist approaches to mobility in China take these arguments seriously (Manion 1985; Burns 1994; Chan 2004; C. Li 2004; Bo 2004, 2010; Kou 2010; Brødsgaard 2012). Work in this tradition places the strategies pursued by the CCP at the center of the analysis and studies the rules, both formal and informal, that govern how vacancies are generated and filled. It finds that the decentralization of the Chinese state in the wake of its transformation from a planned to a market economy (Li and Bachman 1989; Lieberthal 1992; Qian and Weingast 1996, 1997; Landry, Lü, and Duan 2017; Xu 2011) was accompanied by tight control over personnel decisions (Landry, Lü, and Duan 2017; L. C. Li 2010; Naughton and Yang 2004; Xu 2011; Zheng 2007). Modeled after the Soviet nomenklatura system, an elaborate apparatus was put in place that, similar to human resource departments in large firms, manages the process of filling vacant positions (Manion 1985; Burns 1994; Chan 2004) and serves as an important organizational pillar of one-party rule (Edin 2003; Zheng 2010;

McGregor 2010; Brødsgaard 2012).[3]

The remainder of the paper is an attempt to study the strategies pursued by the CCP through an analysis of interorganizational transfers.

VACANCY CHAINS AND ORGANIZATIONAL STRATEGY

So far, we have argued for the usefulness of an organization-centered perspective that takes the organizational strategies behind mobility seriously. But how does one study those strategies? The secrecy with which high-level hiring decisions are made in the Chinese party-state prevents us from observing them directly. As we saw above, scholars in the institutionalist tradition have relied on official documents produced by the CCP such as the nomenklatura lists to study changes in the formal rules that govern appointments (Manion 1985; Burns 1994; Chan 2004; Brødsgaard 2012). More recent work has taken a different approach and has sought to infer informal rules from regularities in appointments (C. Li 2004; Bo 2004, 2010; Kou 2010). We follow this line of work. If we assume patterns of transfers reflect organizational strategies, we can use those patterns to infer strategies. More specifically, the patterns we focus on in this paper are vacancy chains. Vacancy chains allow us to make visible the invisible hand of the Chinese bureaucracy.

Conventional methods for studying mobility assume careers are independent. The moves

3. The *nomenklatura* specifies a list of important positions over which party committees on each administrative level have direct authority. The list covers a wide array of government and non-government positions including the heads of the executive, legislative and judicial branches of the state, the heads of party organs, managers of large state-owned enterprises and banks, principals of government-organized non-governmental organizations (GONGOs) like guilds, workers' unions, and women's confederations, and even editors-in-chief of major newspapers. As a result, "no leader's appointment, promotion, transfer or removal can be effected without the approval (*pizhun*) of the appropriate Party committee" (Manion 1985, 207). In addition, almost all important positions on all administrative levels are held by CCP members. Since 1984, the Central Committee has been officially responsible for appointing provincial and ministerial positions one level down the administrative hierarchy, and provincial party committees have been responsible for appointing positions one level down at the municipal/bureau level (Landry 2008, 46–50). A reform in 1990 following the Tiananmen crackdown gave some local appointment authority back to the center. In practice, the decision-making process is more oblique and usually takes place within a smaller circle of top officials. Often decisions about local appointments are the result of a bargaining process between central and provincial leaders (Burns 1994).

of one person do not affect others. In many mobility systems, however, especially those within organizations rather than whole societies, structural constraints on career moves violate this assumption. White (1970) developed the analysis of vacancy chains to study such highly constrained systems. He argued that in relatively closed systems in which a well-defined set of incumbents moves between a well-defined set of positions and in which each position can be held by only one incumbent, vacancies, not people, have the initiative. His key idea was that, in such a system, mobility depends first and foremost on the availability of vacancies. A person can only move if another position is vacant. And, so, career moves and the relationships between them are features of organizations, not individuals.

Vacancy chains can be observed when positions are filled from within the system. An initial transfer out of the system of positions generates a vacancy in position A, which is filled by someone from position B, creating a new vacancy in position B, and so forth until the vacancy leaves the system of positions under investigation. A vacancy chain is the sequence of moves a vacancy makes from its initial entry into to its final departure from the system. Individuals moving through positions produce careers; vacancies moving through positions produce vacancy chains. From this perspective, careers, the usual focus of sociologists, appear as the end products of a long vacancy-filling process. It is obvious that the moves people make in such a system are highly dependent (Chase 1991, 134).

Following White (1970)'s original study of U.S. clergy, scholars have applied vacancy chain analyses to a range of substantive issues including the mobility of police officers (Stewman 1975b, 1975a), football and basketball coaches (Smith 1983; Smith and Abbott 1983), people in the housing market (Marullo 1985), hospital superintendents (A. Abbott 1990), employees in large firms (Pinfield 1995), and state officials (Yeung 2007).

The bureaucratic nature of the Chinese party-state makes it suitable for a vacancy chain analysis. The vacancy chain model was developed for the mobility of high-status personnel whose entire careers unfold within a single organization (White 1970). Unlike the U.S., where

people can enter high-level political office through lateral transfers from outside the national government (Heclo 1977; Mackenzie 1987), the Chinese state resembles what Putnam (1976) calls a guild system. Ambitious individuals enter the party-state early on in their careers and complete long apprenticeships within the state before entering elite positions. This creates an internal labor market (Doeringer and Piore 1985) with few transfers in and out of the political system, which, in turn, produces a stable population of people competing for a stable population of positions. Throughout our time period, some positions do enter and some do leave the system, but this was a gradual process. It is uncommon that new positions were created ad hoc to accommodate position-less officials.

We argue that vacancy chains allow us to approach mobility from an organization-centered structural perspective. This argument is illustrated in an innovative study by Pinfield (1995). Pinfield was interested in the staffing practices of a large multiplant manufacturing company in the forest products industry. In one of the interviews he conducted with managers of the company, the interviewee describes a situation in which employees were promoted not because they had acquired certain attributes, but merely because a vacant position had to be filled: "We had two openings—one at the pulp mill and one in the sawmill. These positions had been used as training schools for accountants. But once one or two people at the top move, the chain starts and we lose good, experienced people and don't get suitable replacements" (266). The employees in this example were not transferred because of their skills or ambitions, but because vacancies appeared above them in the organizational hierarchy and the management of the firm decided to move them into those vacancies. Rather than actively moving into a higher position, they were moved by the necessities of the system and the decisions of managers.

This example contains the link between vacancy chains and organizational strategies. If we understand vacancy chains as bundles of hiring decisions, we can treat them as residues of the strategies that gave rise to those decisions. We are aware of only two studies that

explicitly link vacancy chains to organizational strategy. The first is Yeung (2007)'s work on the suppression of the Taiping rebellion in nineteenth-century China. Yeung demonstrates that the Chinese imperial court first responded to local turmoil by deploying outsiders or officials from the central administration to replace inept provincial governors. However, the strategy proved to be inefficient in countering the rebels who had better local knowledge and impeded normal vacancy flows in the system. The imperial court eventually changed its recruitment strategy and started to fill vacancies by drawing on closely connected local elites. The strategy proved to be successful in appeasing the rebels, but due to war-time exigencies and the inter-dependency of vacancy replacements, it also created far-flung consequences that cascaded through the entire bureaucratic system.

The second one is Pinfield (1995)'s study mentioned above. Similar to Yeung, Pinfield understands vacancy chains as bundles or strings of hiring decisions that can be used to study organizational strategies. But his study is unique in that he combines his analysis of vacancy chains with interviews with managers. This allows him to link the structure of vacancy chains directly to what the managers say are the decisions that produced them. He finds that short chains were often the response to changes in the local context of a division within the company and the result of local adjustments. In contrast, long chains often occurred in situations in which stable career lines had been established or where hiring decisions were concentrated, such as when one manager was able to coordinate several decisions. Long chains were also used to introduce change. They indicated strategic staffing, that is, new strategic directions led to a series of chained transfers. Thus, what distinguishes long vacancy chains from isolated transfers is their strategic and future-oriented character.

ORGANIZATIONAL STRATEGIES

Different organizational strategies have different implications for the structure of vacancy chains we observe. In particular, we argue that we can use the relationship between vacancy

chains and career success to arbitrate between different strategies. Here we draw on work by organizational scholars and scholars of the state to identify three such strategies and show that they generate different expectations about the relationship between vacancy chains and career success. Below we will confront those expectations with the data to inform our understanding of what guides mobility in the Chinese party-state. Before we discuss these strategies, however, we must concede the possibility that there is no strategy at all or that the strategies that are being pursued by the CCP are not reflected in vacancy chains. In that case, we may expect to find no vacancy chains at all as hiring decisions are not coordinated. Alternatively, we might find vacancy chains simply because they occur by chance. In that case, there would be no systematic relationship between being involved in long vacancy chains and career success. Some people's transfers are embedded in chains and some people's transfers are not, but this does not affect their likelihood to achieve high party ranks.

Promote experience

Organizational scholars have argued that functional specialization and the resulting need for cooperation induce organizations to change their goals from technical to administrative ones (Perrow 1961). Doing so heavily emphasizes the need for people with managerial skills. By rotating employees through different subunits of the organization, they encourage employees to develop generalist skills, to acquire knowledge about how the different organizations operate, and to build social capital—all important resources that allow them to become future organizational leaders. This is the reason why in many organizations there is a positive link between job rotation and career success (Campion, Cheraskin, and Stevens 1994; Ortega 2001; Van Maanen 1984). Further, by rotating people through a variety of different positions, organizations stress-test promising employees and socialize them into the culture of the organization (H. Kaufman 1960; Kanter 1993; Pinfield 1995). One of the

managers interviewed in a study of a large manufacturing company describes it like this: "He'll be rotated through a number of supervised assignments. But he'll also be thrown into the deep end a few times, more of a trial-by-fire experience, to round out his supervisory training" (Pinfield 1995, 300).

Scholars of the Chinese state have noted the same phenomenon. Officials are more likely to advance and reach high-level positions if their careers involve a large number of inter-organizational transfers (Kou 2010; Bo 2010; Jia and Xu 2018). These findings indicate that the Chinese state values diverse job experiences and promotes those who have worked in various organizational units.

If the dominant strategy is one that promotes those with the most diverse experience, we should be able to confirm the findings of previous research and find a positive relationship between the total number of inter-organizational transfers and career success. Beyond that, however, we would not expect an additional effect of being involved in longer vacancy chains. In other words, it could be that being involved in vacancy chains picks up the fact that the official has made transfers. But conditional on the number of transfers, this relationship should disappear.

This line of reasoning, however, is challenged somewhat by the case of the "Shanghai Gang." The Shanghai Gang is perhaps the most notable faction during our time period. After the Tiananmen crackdown in 1989, the CCP's senior leaders chose Jiang Zemin, the party secretary of Shanghai, to fill the power vacuum in Beijing, and Jiang eventually became China's top leader for the next 10 years. To consolidate his power, Jiang promoted many of his former subordinates from Shanghai to fill the country's top positions. As Jiang's power base, Shanghai was historically known to have a tradition of staffing its politicians locally and internally. In fact, some of Jiang Zemin's best-known allies had their entire careers developed in Shanghai before taking top positions in Beijing.[4] Thus, even though those

4. One example is Wu Bangguo, the second highest ranked CCP leader from 2003 to 2013. Wu was never transferred out of Shanghai before entering the Politburo.

individuals were extremely powerful, they had no inter-organizational transfers, implying a negative relationship between the number of transfers and career success.

Sponsorship

The primary goal of organizations is renewal. In order to survive, organizations need to ensure a stable flow of the right people into the right positions (Mosca 1939). As Pinfield (1995, 264) found in his study of a large multi-plant manufacturing company, "[m]ost managers recognized that they needed to keep the human resource system of the corporation moving with a steady stream of staff into, through, and out of the organization." To accomplish this, organizations have built human resource departments that are tasked with hiring, training, and mentoring employees. And they are guided by a vast management literature that addresses issues such as "staffing" and "succession planning."

Often staffing is not simply about filling vacant positions. Companies view hiring decisions with an eye towards the future development of the candidates as well as that of the organization. Pinfield found that advancement potential or promotability was a crucial criterion when selecting candidates for a vacancy and the selection was followed up with a long-term process of monitoring, assessing, teaching, and coaching. Job transfers were part of a larger strategy of grooming employees and paving their way towards leadership positions in the organization.

Li and Walder (2001) have suggested that the positive relationship between party membership and career advancement that had been found in previous studies should be understood as the result of a strategy of sponsorship. According to the authors, the CCP actively incorporates promising youth into the party early on in their careers and subsequently grooms them by giving them access to a range of career advantages.

As a modern Leninist party, the CCP has always been highly conscious of organizational renewal and the selection of its "Communist heirs." During the early years of the PRC, it

selectively sponsored people who belonged to certain "class origins" to maintain the ideological purity of the party-state (Treiman and Walder 2019; Xie and Zhang 2019). After 1978 Deng Xiaoping and the CCP's senior leaders soon sensed the need of bringing in younger and more educated cadre and wrote cadre "juvenilization" (*nianqinghua*) and "professionalization" (*zhuanyehua*) into the Party Constitution.[5]

Sponsorship through special career tracks has an important consequence. Job security, in conjunction with predictable upward mobility, produces career orientations which, in turn, help align the interests of employees with those of the organization (Weber 1958). As long as an employee sticks with her firm, she is likely to move up the status hierarchy without having to worry too much about arbitrary dismissal or demotion. This creates loyalty and encourages employees to identify with their organization (Blau 1955). In that sense, sponsorship can be understood as a form of social control in which career opportunities are exchanged for loyalty.

The rapid growth of the French military in the sixteenth and seventeenth century and the concomitant expansion of the state's administrative apparatus to finance and manage it, for example, created job opportunities for notables through state employment, which, in turn, appeased them and created political stability (Duffy 1980). Another noted example is Stalin's Great Purge (Padgett 2012, 272). Although it greatly weakened the power of Soviet old elites, the vacancies it generated put a whole generation of "red engineers" into fast tracks and paved the way to the Soviet Union's gigantic output in heavy industry. In China, careers were also used as a way to draw local elites into the orbit of the central state and to align their interests with those of the state (Kuhn 1980). Once elites are tied to the spoils of the central administration, transfers can be used as leverage against bureaucrats whose long-term career ambitions are linked to the prestige of their posts (Iyer and Mani

5. Deng first proposed the idea of "juvenilization," "intellectualization" (*zhishihua*), and "professionalization" in a talk called "The Reform of the Party and State's Leader System" (*dang he guojia lingdao zhidu de gaige*) on August 18, 1980. The terms were codified at the 12th Party Congress held in September, 1982.

2012). Stalling those political ambitions, on the other hand, can quickly result in political disintegration and even rebellion (Gould 1996).

What if vacancy chains were the residue of such a strategy of organizational sponsorship? As we said above, the difference between a long chain and an isolated transfer is the former's strategic and future-oriented character. A single transfer is ad hoc; a long chain is like a well-planed sequence of moves in a chess game that carefully develops each piece so that it can realize its full potential. Under a strategy of sponsorship, we would likely see a positive relationship between being involved in vacancy chains and career success. In addition, we would expect that many people are involved in vacancy chains relatively early in their career. Being part of a vacancy chain from early in a career signals an official has been identified as promising and is incorporated into a sponsorship system.

Control

Alternatively, instead of promoting promising officials, the Chinese state might pursue a strategy of controlling people who cause problems. Organizational scholars in economics and sociology have long understood that the issue of control is at the heart of organizations (Moe 1984; Perrow 1986; Fligstein 1990). Often the interests of an organization and the interests of its staff do not automatically align and the organization is unable to fully regulate the actions of its staff through contracts (Jensen and Meckling 1976; Fama and Jensen 1983). Even with good intentions, organizational men often develop a vested interest that serves their own needs rather than the manifest goals of the organization. Berliner (1957), for example, documents how factory managers in the Soviet Union had self-interest to under-report their production capacity and over-report their production output. Sending inspectors to local units was costly and therefore infrequent, and whenever an inspector became permanently stationed in a local unit, his interest became soon tied to those of the local organization (see also Devons 1950).

The problem of control is amplified as organizations grow larger and differentiate along functional or geographic lines (Blau 1970). As private companies significantly increased in size and began to offer increasingly complex products at the end of the nineteenth century, they adopted a multidivisional organizational form in which a general office at the top was responsible for strategic planning and distinct product-based or region-based divisions below were responsible for tactical planning and execution (Chandler 1962). These structures made the direct supervision of all employees impossible and amplified the information asymmetries between central management and divisions. These organizational structures are not restricted to private companies but are also common in state bureaucracies, including the Chinese party-state (Carpenter 2001; Xu 2011; Qian and Weingast 1996, 1997).

Organizations countered the centrifugal forces brought about by growth, specialization, and differentiation by developing a range of control mechanisms to ensure that individuals and subunits within the organization would act in the interest of the organization as a whole (Doeringer and Piore 1979; Kellogg, Valentine, and Christin 2020; Simon 1947; March and Simon 1958; Freeland 1996; Eccles and White 1988). Rotating people between subunits of the organization is one such mechanism of control. Transfers are a way to weaken centrifugal and strengthen centripetal forces within large organizations. A study of the U.S. Forest Service, for example, shows that forest rangers are regularly transferred from one station to another to prevent the formation of strong ties to local communities that made them more responsive to the interests of local stakeholders than those of the agency (H. Kaufman 1960). By cutting ties to outside interests, the agency also strengthened the rangers' identification with the organization. "Only one thing gives any continuity, any structure, to [the forest ranger's] otherwise fluid world: the Service" (178). Finally, office rotation made it difficult for individual rangers to hide practices that were not compatible with the agency's goals. While a ranger might have been able to successfully hide her practices from her superiors, she could not hide them from her successor.

States, too, grapple with the issue of control, which arises from their centralizing tendencies. States centralize political authority, leading to tensions between the national center and local interests (Weber 1958; Poggi 1978; Tilly 1985, 1992; Barkey 1994; Mann 2012). In regimes that govern over large territories, centrifugal forces are particularly strong. Under these conditions, the center faces the "risk of disloyalty, dissimulation, corruption, and rebellion" (Tilly 1992, 25). This can be seen in the Chinese case. China's history is full of conflicts between the central administration and warring local elites and rebellious peasants (Mousnier 1970; Kuhn 1980; Perry 1980).

Work on the Ottoman Empire (Barkey 1994), for example, documents that rulers often use office rotation among localities as a means to prevent officials from establishing strong ties to their local communities and from becoming too autonomous. Office rotation, thus, compensates for the center's lack of insight into local conditions. The same was true in Imperial China (Ho 1962). According to the rule of avoidance, officials were prohibited from being assigned to positions in their home region. Often these transfers were not simply enforced by the central government but were the result of negotiations between central and local elites. Chen, Wang, and Zhang (2021) recently found that as early as the 400s the Northern Wei Dynasty started to offer greater upward mobility to elite families living in regions with paramilitary strongholds in exchange for their willingness to take public offices outside their power base. A slightly different strategy was pursued by the Manchu regime during the Taiping Rebellion when officials with close links to the metropolitan center were deployed to rebellion-ridden provinces to control local activities (Yeung 2007; Zhang 2021).[6]

Many have argued that the reshuffling of officials—by moving them from province to province, from center to province, or from province to center—is still an important control

6. Japanese rulers in the fifteenth and sixteenth century pursued a somewhat different control strategy based on the physical movement of people, but not necessarily through appointment to positions. According to a policy called *sankin-kōtai*, feudal lords were required to reside alternately in the ruler's castle and their own and, when absent from the center, to leave their families behind. This strengthened central control over local elites.

strategy in modern China (Dittmer 1978; Walder 1995a; Huang 2002; Bo 2004; Zeng 2016; Eaton and Kostka 2014; Wang 2022) In fact, since the beginning of the 1978 reforms, which have dramatically decentralized the state apparatus, control efforts have likely increased (C. Li 2004; Brødsgaard 2012).

If vacancy chains are the product of a strategy that primarily tries to keep (potentially) threatening officials down, we might expect a negative relationship between being involved in vacancy chains and career success. Being part of a chain could signal that an elite has been identified as a (potential) threat to the organization. This would mean that compared to people whose transfers are isolated, officials in vacancy chains fare worse and that they would have done better had they not been part of a chain.

There is another reason why we may expect a negative relationship between being involved in vacancy chains and the highest rank reached in the party. Organizational units within state bureaucracies tend to seek some degree of autonomy and one of the most important types of administrative power they would like control over is the ability to choose their staff. When a vacancy is filled with an external hire rather than an internal one, this could indicate the weakness of an organizational unit. In other words, if we follow exclusively the control logic, it might be the case that cadres and organizational units involved in long chains are the ones that the central state is able to control. The organizational units have less ability to bargain with the central state and shield their local interests. Officials in those units have less ability to plan and control their destinations. In contrast, elites in more powerful organizational units can move into desirable openings in other parts of the system, and the vacancies they produce are not passed on to other organizations but filled internally. Burns (1994, 471) provides an example of such a scenario in the case of Guangdong province. From 1989 to 1991, when the central authority tried to replace Ye Xuanping, the Governor of Guangdong, by promoting him to the center, Ye reportedly resisted the transfer until he was able to negotiate a deal with the Center and put his own man from within the province

into the position he left. If it is true that individuals in long chains have less political power, we might also expect them to have a lower chance of reaching high party ranks.

On the other hand, one might argue that only powerful, and thus high-ranking, officials are exposed to this type of control through office rotation. In that case, we would expect a positive relationship between being involved in vacancy chains and party rank. But we would then also expect that people's transfers appear in vacancy chains close to the height of their career when they have advanced enough to pose a serious threat. In addition, it is likely that vacancy chains would tend to move horizontally, replacing people of similar rank.

DATA AND METHODS

We draw on a novel dataset of inter-organizational transfers of Chinese officials covering the period from 1977 to 2012. Data on the job transfers of Chinese officials is not readily available. We apply computer-assisted coding to the CVs of a comprehensive list of all high-level Chinese political elites contained in the Chinese Political Elite Database, maintained by National Chengchi University (NCCU) in Taiwan.

In CCP studies, specialists usually start by defining the level of elites they want to include and then collect and hand-code as much public information as possible about the selected elites. For example, Shih, Adolph, and Liu (2012) hand-coded the work histories of all Central Committee and Alternate Central Committee members from 1982 to 2007. Jiang (2018) maintains a dataset that covers all provincial party standing committee members and city party secretaries and mayors from 2000 to 2012. And a team led by Xueguang Zhou (Zhou et al. 2018; Zhou et al. 2021) maintains a dataset that covers the complete career histories of nearly all bureaucrats in Jiangsu Province.

The NCCU database we draw on is by far the most comprehensive data source of high-level CCP elites. According to its website, it contains the CVs of all Chinese political elites who reached the level of vice-governor or vice-minister or above since 1990 or who reached

the level of governor or minister or above since 1966. We include all CCP elites who were born after 1923[7] and active within our study period (N = 4,290). This gives us the ability to obtain a comprehensive picture of the Chinese bureaucracy during China's reform period.

The database provides text descriptions of all jobs an elite held during his or her career. Based on these descriptions, we were able to code the organizational unit of each job. The CCP relies on a *tiaokuai* system to administrate the vast territory of the party-state. This system is composed of various functional ministries (*tiao*) and around 30 provincial units (*kuai*). Each provincial/ministerial unit acts as an independent organization and oversees a geographical/functional domain. We coded all provincial and ministerial units, broadly defined. This includes not only all provinces, municipalities, autonomous regions, and ministries under the state cabinet but also party organizations, non-government organizations, as well as many institutions that enjoy the rank of a ministry or vice-ministry. Although functional agencies also exist at the provincial level, not just the national level, for simplicity, we only coded ministries in Beijing. We coded all positions at the provincial level as the province in which they occurred. We excluded the People's Liberation Army (PLA) from this study because job transfers between civilian posts and military posts are almost nonexistent in the reform era (Jia and Xu 2018). We also did not code honorary positions and part-time jobs. A list of all 249 organization can be found in Appendix A.1. Job transfers and vacancy chains were identified based on these organizations. For some of the analyses, it will be useful to group organizations instead of treating them individually, which is why we assigned each organization to one of four domains: ministry, province, party, and other.

Every time a vacancy opens up in one of the 249 organizational units, a decision needs to be made as to how to fill that vacancy. Generally speaking, a vacancy can either be filled internally or through a transfer from a different organizational unit (Carlson 1961; Grusky

7. The cutoff year was chosen such that people included in this study were at most 25 years old in 1949 when the PRC was founded. They spent the majority (if not all) of their career after the civil war and had few revolutionary credentials. In our descriptive analysis, however, we included all elites who were active during the period to provide a global analysis of all transfers and vacancy chains.

1964). Internal hires are likely to prioritize area-specific expertise, and external hires prioritize generalist skills and facilitate communication and integration between organizations. As such the distinction between internal and external hires is related to the Weberian distinction between bureaucrats, whose careers are characterized by career advancement within an organization, and politicians, whose careers tend to include lateral moves across organizations (Aberbach, Putnam, and Rockman 1981).

Our dataset contains only inter-organizational transfers. The complexity of the text source made it too costly to code an official's exact position within an organizational unit, making it impossible to trace movements within units. But the focus on inter-organizational transfers is also justified by the fact that it is these types of transfers that form important steps in the careers of high-level Chinese officials (Kou 2010; Bo 2010; Jia and Xu 2018). And we observe an inter-organizational transfer when an external hire is made. As a consequence, vacancy chains occur in our data when multiple external hires are chained together. A vacancy chain terminates in our data when it is filled internally, that is, by promoting officials within the same organization, when it leaves the system of positions under investigation, or when it remains vacant indefinitely.

For each inter-organizational transfer in our dataset, we searched for potential subsequent moves which, within a one-year window,[8] ended in the organizational unit where the first move departed. We used a fuzzy-matching algorithm to narrow down the list to jobs that had similar text descriptions to the first one. We employed four human coders (including the first author) to go through each entry and pick the exact successor. Eighty percent of all entries were examined independently by at least two coders, including all cases that the first coder was uncertain about. After linking all job transfers into vacancy chains, we manually

8. The one-year window (six months before and six months after the transfer) is an arbitrary choice. A larger window would have made it difficult to determine whether the new person really was a successor of the departing person. But more importantly, if a vacant position is filled too long after it was vacated, it is hard to imagine that the two transfers are part of the same strategy. Also note that the time window only helped us identify possible candidates for successors; each candidate was then checked manually.

searched through the CCP's official documents and identified the predecessor of the first elite in the chain (the person who initiated the vacancy chain) as well as the successors of the last elite in the chain (that is, the person who terminated the chain).[9] We obtained a total of 2,528 transfers, which occurred during the study period. 781 of those appeared in 324 vacancy chains; 1,747 were isolated transfers. Our coding procedure helped us examine the completeness of the CPED database. In the final version, no chain is incomplete due to elites not being included in the database.[10]

To examine the relationship between vacancy chains and career success, we first estimate person-level OLS regressions. To measure career success, we followed standard practice and obtained a measure of the maximum party rank an official reached in his or her career (*Max Rank*). Party rank was coded as follows, lowest to highest rank: 1 = Alternate Central Committee (ACC), 2 = Central Committee (CC), 3 = Politburo (PB), 4 = Politburo Standing Committee (PBSC). If a person was not in any of the committees of a Party Congress, his or her rank is 0. The careers of people who were born in recent cohorts had not ended by the time this research was conducted, and hence their maximum ranks are unobserved. We therefore only included people born before 1954.[11] As a consequence, the number of elites included in the person-regression is reduced to 2,797.

9. 42.4% of the vacancy chains were initiated by retirement. Other initial vacancies occurred because a new position was created (17.3%), the predecessor became a national leader (12.9%), was internally transferred to another position inside the same organizational unit (10.9%), externally transferred out of the system (7.5%), or unexpectedly ended the job due to either death or prosecution (5.5%). 60.2% of the chains terminated as a result of recruitment from within the same organizational unit, 11.5% because of recruitment from outside the system, and 11.0% remained vacant for at least 6 months. For the remaining 17.3% of the cases, we were unable to find the exact records of their replacements. However, we at least made sure that no individual in the database replaced the positions. Those positions are at relatively lower ranks and most likely filled with relatively less known people from within the organizations.

10. Because we tracked every chain from start to end, the only way a chain could be incomplete is if it involves an individual who is not included in the CPED database. During the process, we found 10 party officials who were involved in vacancy chains but were not included in the CPED database. We manually created new data entries for those individuals to make all chains complete.

11. Our data collection ended in 2012. People born between 1950 and 1954 still had one more chance of getting into the Politburo at the 19th Party Congress held in October 2017. Thus, a more conservative approach would have been to only include people born before 1950. However, very few of the 1950-54 cohort still had this chance. The impact of not observing the 19th congress is probably negligible.

To test the importance of diverse job experience, we include the number of inter-organizational transfers officials make (*Transfers*). As discussed above, previous work on the Chinese bureaucracy has shown that the number of transfers is positively related to career success. Our task, then, is to test whether there is an effect of being involved in longer chains even after controlling for the total number of transfers. Because entering the Politburo often induces certain transfers, only the number of transfers up to that point are counted.[12]

To measure the role of vacancy chains, we utilize the distinction between isolated transfers and transfers that are embedded in longer vacancy chains. In the case of an isolated transfer, a person moves into an initial vacancy and the position he or she vacates is where the vacancy chain terminates. In other words, an isolated transfer is a vacancy chain of length 1. In the case of longer vacancy chains, an initial vacancy produces a cascade of transfers. Figure 1.1 illustrates this distinction. It shows two vacancy chains from our data. In the first chain, an initial vacancy occurs in the State Commission Office for Public Sector Reform (*zhongyang jigou bianzhi weiyuanhui bangongshi*) due to the retirement of its standing deputy director. This vacancy then migrates through a set of other organizations until it reaches the newspaper *People's Daily* where it is internally absorbed. The chain has a length of five as it includes five inter-organization transfers. These five transfers were made by five different people, all of whom are therefore involved in a vacancy chain of length five. The second chain originates in the Central Propaganda Department, moves to Zhejiang and is immediately absorbed. It has a length of one, which we characterize as an isolated transfer. Note that the fourth transfer in the first chain and the transfer in the second chain are similar as they both have the same destination (the Central Propaganda Department). Only one of them is embedded in a long chain, however, the other is not. More specifically, we use two variables, the number of times an official's transfers were part of longer vacancy chains

12. Note that our data include officials with no transfers.

instead of appearing in isolation from other transfers (*Chained Transfers*) and the average length of the vacancy chains an official was involved in (*Average VC Length*).

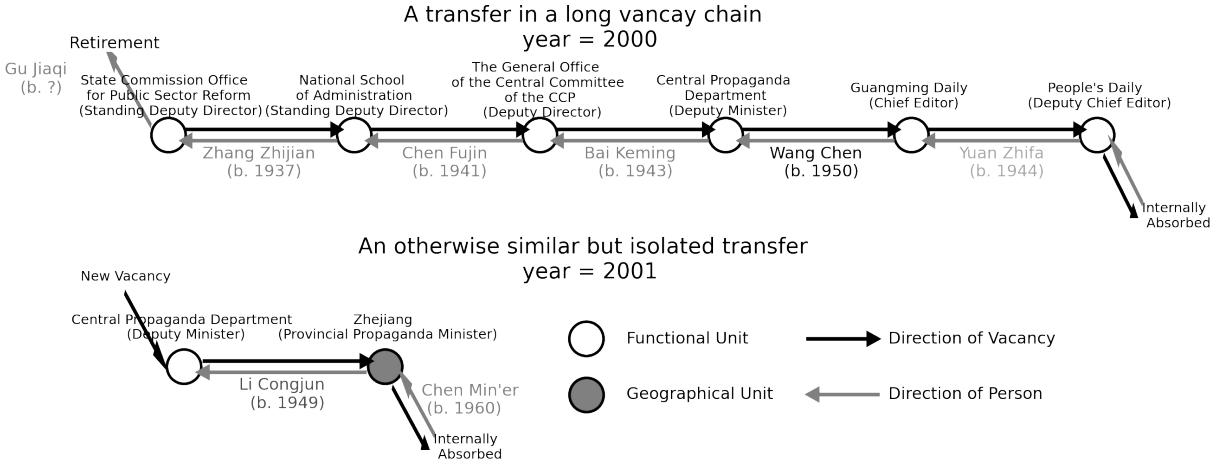

Figure 1.1: A long vacancy chain (length = 5) and an isolated transfer (length = 1)

Basic demographic variables are included such as birth year, gender, whether the person belongs to an ethnic minority, his or her professional title, and whether he or she studied abroad.[13] Controlling for these variables allows us to see how much the patterns we observe can be explained by individual-level characteristics.

In a second analysis, we switch from the person-level to the transfer-level. We use a binary dependent variable that measures whether the person was promoted at the next Party Congress or not (*Promotion*). Our main predictor is the length of the vacancy chain each transfer was embedded in (*VC Length*), ranging from 1 to 7. Switching to the transfer-level allows us to do several things. First, we can now control for the party rank at the time of the transfer to ensure that our model captures the effect of being involved in a vacancy

13. We created two additional variables that we decided to omit from the models: an indicator capturing whether the person has a college degree and the age at which the person joined the CCP. Because our sample includes relatively high-level elites, only 221 people have no college degree, and those who make it to the top without a college degree are very special people (such as Xie Fei, the Provincial Party Secretary of Guangdong from 1991 to 1998, who joined the party at the age of 17), which is why we decided not to include this variable. We also do not include the age when entering the party because its effect turns out not to be significant. In addition, due to missing data, including these variables would have resulted in fewer observations.

chain and not prior status differences (*Year Rank*). Second, we can include an interaction effect between *VC Length* and the *Age* of the person at the time of the transfer to see if there is effect heterogeneity across people of different ages. Third, transfer-level analyses also allow us to control for the *Destination* of each transfer. We do this by conducting more targeted comparisons between transfers that end in the same domain. Because some transfers are made by the same person, the units of analysis in the transfer-level models cannot be treated as independent. We estimate hierarchical logistic regressions with transfers nested in persons to obtain standard errors that are adjusted appropriately (Raudenbush and Bryk 2002).

ANALYSIS

An example

Before examining the relationship between vacancy chains and career success, it is useful to discuss one example vacancy chain in more detail and to present some general observations about the structure of vacancy chains in our data. As for the example, we can return to the vacancy chain portrayed in Figure 1.1 (top). It shows the way in which vacancies cascade through the system and how a long vacancy chain can generate opportunities for younger politicians further down the administrative hierarchy. The vacancy is one that circulates almost entirely within the party domain. The vacancy chain starts in the State Commission Office for Public Sector Reform, the executive organ of a commission under the CCP's Central Commission, when the position of deputy-director was vacated. The commission is responsible for making policies related to administrative reform. Although the position enjoys the rank of a minister, it is not a particularly prominent one, which is reflected in the fact that the departing deputy-director was not a CC member. Zhang Zhijian, the standing deputy-director of the National School of Administration (who also enjoys the rank of a

minister) took the first vacancy and left his position to be filled. Zhang was already 62 years old at the time and was near the retirement age of 65. Unless he could make it into the Politburo in the next Party Congress, this was likely going to be the last job in his career. Because his position was already high, and he was already a CC member at the time, his chance or receiving a further promotion was slim. The vacancy was then passed on to Chen Fujin (age 58) and then to Bai Keming (aged 56), both of whom were deputy-ministers and not CC members at the time but could still enter the CC in the next Party Congress. Then, Bai Keming's deputy-minister position in the Central Propaganda Department was filled by Wang Chen, the chief editor of *Guangming Daily*. Guangming Daily is a sub-ministry-level party organ that is under the supervision of the Central Propaganda Department. At the time of the transfer, Wang, whose entire career until then had unfolded at Guangming Daily, was at the age of 49. He still had a long career ahead of him. The deputy-director position acted as a steppingstone. After that, he served as the head of two ministry-level units and entered the Politburo in the 19th Party Congress. The last transfer from People's Daily to Guangming Daily was more like an ad hoc transfer that filled the vacancy Wang left as his successor was more than 6 years older than him and did not have much room to advance further.

Descriptives

The starting year of our study corresponds to the beginning of the CCP's economic reform. What followed was a gradual institutional departure from the socialist central planning system that the CCP had imported from the Soviet Union (Shirk 1993; Naughton 1995). Many central planning agencies became obsolete, ministries were downsized or merged with others, factories and state-owned enterprises were decoupled from ministries and began to operate as standalone firms (Burns 1993, 349; Lan 2000). At the same time, fiscal decentralization, the establishment of special economic zones, and regional experimentation led to

more autonomy for regional officials.

These dramatic changes, which unfolded over the course of three decades, are picked up by our data. As the system no longer operated based on a centralized command structure, it started to rely more heavily on elite shuffling as a means for integration. As a result, longer vacancy chains became more frequent over the course of our study period, and especially after 1992 when the CCP officially legitimized the market economy and further accelerated the reform (Table 1.1).

Table 1.1: Length distribution of vacancy chains by time period

	1	2	3	4	5+	All
1977-1981	204	11	2	1	0	218
	(93.6%)	(5.0%)	(0.9%)	(0.5%)	(0.0%)	
1982-1986	237	10	4	0	0	251
	(94.4%)	(4.0%)	(1.6%)	(0.0%)	(0.0%)	
1987-1991	271	18.5	9.5	0	0	299
	(90.6%)	(6.2%)	(3.2%)	(0.0%)	(0.0%)	
1992-1996	292.6	28.4	10.5	0.5	1	333
	(87.9%)	(8.5%)	(3.2%)	(0.2%)	(0.3%)	
1997-2001	336.5	40.5	14.5	5.5	2.5	399.5
	(84.2%)	(10.1%)	(3.6%)	(1.4%)	(0.6%)	
2002-2006	251	30	16	5	3	305
	(82.3%)	(9.8%)	(5.2%)	(1.6%)	(1.0%)	
2007-2011	163	45	28	7	8	251
	(64.9%)	(17.9%)	(11.2%)	(2.8%)	(3.2%)	
All	1755.1	183.4	84.5	19	14.5	2056.5
	(85.3%)	(8.9%)	(4.1%)	(0.9%)	(0.7%)	

But not only did long vacancy chains become more frequent, they also experienced fundamental changes in how they moved through the system of positions. This is shown in Tables 1.2 and 1.3, which give estimates of opening and terminal probabilities by time period and organizational domain.[14] The opening probability is the probability, for each domain, that

14. To be able to calculate opening and terminal probabilities, we need to impose some structure onto the data. We therefore group the original 249 organizations into four domains and calculate probabilities for each domain.

a long vacancy chain (of length 2 or above) originates in an organization belonging to that domain. For example, Table 1.2 shows that 48% of the long chains that occurred in the time period from 1997 to 2001 originated in provinces, while only 19% originated in ministerial organizations. Because demotions are rare in the Chinese system,[15] vacancy chains mostly flow from high-ranking positions to low-ranking positions. Therefore, domains involved at the beginning of a vacancy chain can be considered as more important than those involved towards the end of a chain. The terminal probability is the probability that a long vacancy chain is internally absorbed. Domains that have high terminal probabilities are the ones that are more likely to fill their vacancies with internal hires. It can be seen that in earlier periods, ministries were more likely to generate initial vacancies while provinces were places where vacancy chains terminated with high probabilities. Vacancies moved from the center into provinces as officials moved in the opposite direction. An interesting shift took place around the 15th Party Congress in 1997 when provinces became the domain in which initial openings occurred. At the same time, vacancy chains became much less likely to terminate in provinces. 1998 marks an important year in the CCP's institutional reform. The State Council underwent a major reform. Cabinet ministries were downsized from 40 to 29; all industry-specific ministries, such as the Ministry of Mechanical Industry and Ministry of Forestry, were disbanded or merged to allow for the separation of government agencies and firms; and large-scale layoffs in the state sector took place. The results in Tables 1.2 and 1.3 demonstrate a shift in how the state coordinated the career flows of its most important elite members following those changes. As ministries decreased in importance, the state shifted its focus to provinces. These findings inspire confidence that our vacancy chain analysis is picking up changes in organizational strategy.

15. For all vacancy chains in our data (but not for isolated transfers), we coded the administrative rank of the positions involved in the transfers. Administrative ranks are the ranks of the positions in China's civil service system. Most of the inter-organization transfers took place at the provincial/ministerial, sub-provincial/sub-ministerial, or bureau director levels. Using these ranks, we find that 63% of the transfers are lateral, 31% are promotions, and only 6% are demotions.

Table 1.2: Estimates of opening probabilities by time period and domain

	Province	Ministry	Party	Others
1977-1981	4	2	1	1
	(50.0%)	(25.0%)	(12.5%)	(12.5%)
1982-1986	5	5	4	0
	(35.7%)	(35.7%)	(28.6%)	(0.0%)
1987-1991	9	11	6	0
	(34.6%)	(42.3%)	(23.1%)	(0.0%)
1992-1996	7	24	9	12
	(13.5%)	(46.2%)	(17.3%)	(23.1%)
1997-2001	30	11.5	15.5	5
	(48.4%)	(18.5%)	(25.0%)	(8.1%)
2002-2006	18.5	18	10	5
	(35.9%)	(35.0%)	(19.4%)	(9.7%)
2007-2011	37	21	11.5	10
	(46.5%)	(26.4%)	(14.5%)	(12.6%)

Table 1.3: Estimates of terminal probabilities by time period and domain

	Province	Ministry	Party	Others
1977-1981	6	3	–	–
	(66.7%)	(75.0%)	(–%)	(–%)
1982-1986	7	5	1	–
	(77.8%)	(100.0%)	(33.3%)	(–%)
1987-1991	16	9	1	–
	(80.0%)	(64.3%)	(100.0%)	(–%)
1992-1996	35	10	3	2
	(92.1%)	(57.1%)	(42.9%)	(100.0%)
1997-2001	37	11.5	9.5	1
	(60.7%)	(69.7%)	(63.3%)	(20.0%)
2002-2006	40	9	3.5	2
	(68.4%)	(46.2%)	(33.3%)	(50.0%)
2007-2011	44	12	14	8
	(51.2%)	(36.4%)	(66.7%)	(80.0%)

Vacancy chains and career success

We now turn to results from regressions that examine the relationship between being involved in vacancy chains and career success. We first look at the bivariate relationship between the number of transfers an elite has made and status attainment as measured by maximum party rank. Above we saw that career experiences in different types of organizational units help officials to develop the skills and resources that becoming organizational leaders demands from them, and that therefore inter-organizational transfers are positively related to career success. This is confirmed by our data which suggest that more diverse experience leads to higher party ranks. Model 1 in Table 1.4 shows that each additional inter-organizational transfer increases the party rank by .34, or .40 standard deviations. The effect is highly significant.

Given that transfers are positively related to career success, we can now ask whether being involved in long vacancy chains versus isolated transfers has an additional effect on career success. As discussed above, there are three ways to think about this. First, it is possible that conditional on the number of transfers, there are no systematic differences between people who are part of longer chains and people who make isolated transfers. What matters is how often people transfer, not whether those transfers are part of vacancy chains. Second, we might expect that long vacancy chains reflect a strategy of organizational sponsorship. In this case, we would expect a positive relationship between being part of vacancy chains and career success. Third, one might expect that long vacancy chains reflect a control strategy that tries to keep up-and-coming officials down. If that is the case, and if vacancy chains are an expression of this strategy, we might expect to see a negative relationship between being part of vacancy chains and career success.

To test these conflicting expectations, in model 2, we add the number of times an official's transfers were part of longer vacancy chains and the average length of an official's vacancy chains. The results are unambiguous. Being part of longer vacancy chains is positively related

Table 1.4: Person-level OLS regressions predicting maximum rank

	model (1)	model (2)	model (3)
Transfers	0.34***	0.20***	0.18***
	(0.01)	(0.02)	(0.02)
Chained Transfers		0.32***	0.33***
		(0.05)	(0.05)
Average VC Length		0.05*	0.06***
		(0.02)	(0.02)
Birth Year			-0.03***
			(0.00)
Female			0.03
			(0.05)
Ethnic Minority			0.18***
			(0.05)
Profession (=none)			
economic analyst			0.16***
			(0.06)
engineer			0.30***
			(0.04)
others			0.19***
			(0.06)
professor / research fellow			0.14**
			(0.06)
Studying Abroad (=none)			
Soviet Union / Eastern Europe			0.34***
			(0.09)
other countries			0.22***
			(0.07)
Intercept	0.28***	0.28***	0.18***
	(0.02)	(0.02)	(0.02)
N	2797	2797	2633
R^2	0.16	0.20	0.29
Adjusted R^2	0.16	0.19	0.29
Bayesian Inf. Crit.	6632.89	6537.69	5998.52

Note: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

to status attainment in the Chinese party-state. Each additional time an elite is involved in a vacancy chain increases his or her party rank by .25. On top of that, a one-unit increase in the average chain length is associated with a .08 increase in party rank. Both effects are highly significant. These findings hold even after controlling for individual-level characteristics like birth year, gender, ethnic minority, profession, and educational experience (model 3). While the coefficients for all control variables, except gender, are significant, we note that the R^2 in the final model (model 3) is three times larger than the one in a model that only includes the controls (results not shown). This means that transfer-based measures are much more consequential than individual-level factors for mobility among high-level political elites in China. These findings, then, provide evidence that in addition to promoting officials with diverse experience, the CCP pursues a strategy of organizational sponsorship.

According to our data, Shanghai is one of the few exceptional provincial units that defies the logic of vacancy chains. From 1977 to 2011, Shanghai was involved in vacancy chains only 12 times (1.1% of all units involved in vacancy chains), while it was involved in isolated transfers 55 times (1.6% of all units involved in isolated transfers). Although we saw above that provincial positions played an important role in vacancy chains, Shanghai was 0.71 times as likely to be involved in vacancy chains rather than isolated transfers. On the other end of the spectrum, Xinjiang was 2.75 times as likely to be involved in vacancy chains, followed by Chongqing (2.42 times), Anhui (1.92 times), Yunnan (1.86 times), Tibet (1.81 times) and Inner Mongolia (1.76 times), which are all inland provincial units. In fact, the vast majority of provincial units (26 out of 31) had a higher chance of being involved in chained transfers than isolated transfers. (The full statistics of all provinces can be found in Appendix A.2.) In 64% of the isolated transfers in which Shanghai was involved, it appeared as a sender of the person (the receiver of the vacancy). In other words, elites from Shanghai were more likely to fill ad hoc vacancies, and Shanghai did not pass on those vacancies to other units. Thus, because of Shanghai's powerful position and its tendency to absorb

vacancies internally, strategic moves that led to vacancy chains evaded it. This observation invites further investigation.

A well-known example of sponsorship is the case of Hu Chunhua (see Table 1.5). Hu was born in a farmer's family in a poor inland village and entered Peking University at the age of 16 due to his excellent performance in the college entrance exam. After graduation, he volunteered to work in Tibet and joined the regional branch of the Communist Youth League (CYL). There, his career began to skyrocket. He held various positions in the region and was transferred back and forth between Tibet and CYL's headquarter in Beijing. At the age of 43, he became the First Secretary of the CYL (which enjoys the rank of a minister). Two years later, Hu took the position of the governor of Hebei and became China's youngest provincial leader. As shown in Table 1.5, Hu was involved in a total of six inter-organization transfers in his career, most of which were part of longer chains. Even his very first transfer was part of a chain even though he began his career in a region that is the furthest away from the capital city. As the last person in a three-transfer chain, the vacancy chain "connected" him to two other officials. One of them was a CC member at the time, the other would later become one. Hu's transfer from the CYL to Hebei at age 45 was the fifth transfer of a vacancy chain of length seven, one of the longest chains in our dataset. All cadres who preceded him in the chain were members of the CC. In 2012, Hu entered the Politburo at the age of 49 and became one of the most promising stars in Chinese politics.

That the CYL figured so prominently in Hu Chunhua's career is not an accident. During our time period, many high-profile Chinese politicians had early careers in the CYL. The CYL was founded in the 1920s as a party organ and is by design the party's "reserve army." Its institutional role is to help the government manage youth affairs and also to recruit and train the CCP's future leaders. Its leadership positions can only be held by party cadres who are under a certain age limit. These cadres usually only stay in the organization for a brief period of time before they are ready to take other positions in the party state. Therefore,

Table 1.5: Career of Hu Chunhua

origin	destination	year	chain length	position in the chain	party rank	admin rank 1[a]	admin rank 2[a]
Tibet	CYL	1997	2.5[b]	2.5	None	bureau-director	bureau-director
CYL	Tibet	2001	1.5	1	None	sub-pro/min	sub-pro/min
Tibet	CYL	2006	4	3	None	sub-pro/min	pro/min
CYL	Hebei	2008	5	3	CC	pro/min	pro/min
Hebei	Inner Mongolia	2009	2	1	CC	pro/min	pro/min
Inner Mongolia	Guangdong	2012	2	1	CC	pro/min	pro/min

Note: [a] pro/min stands for the provincial-ministerial level; sub-pro/min stands for the sub-provincial (sub-ministerial) level. [b] Two secretaries of the CYL Central Secretariate left their positions around the same time. Hu Chunhua was one of the two successors. His two predecessors were involved in two chains of lengths 2 and 3 respectively. Because it is difficult to determine who replaced whom exactly, we coded the chain length as the average of the two chain lengths.

the CYL acts as a hub or "stepping stone" in the CCP's career mobility network. It serves as an institutional channel for promoting young and promising elites (Doyon 2019; Kou 2014; Kou and Tsai 2014).

These results provide evidence that a pure control strategy is not what we are seeing. But this relies on the assumption that control implies a negative relationship between being involved in vacancy chains and career success. It is, however, not implausible to argue that only relatively powerful people, who can pose a real threat to CCP elites at the very top, are targeted by these types of control efforts. If powerful people are selected into longer chains, we may expect the relationship to be positive. We therefore need more than the previous analysis if we want to be able to arbitrate between control and sponsorship.

Two additional analyses lend credence to the sponsorship argument. The first one utilizes the difference between people's party rank at the next Party Congress and their party rank at the time when they were involved in transfers. Given a strategy of control that targets powerful people who pose a threat, we should see that elites have already reached or come close to their highest party rank when they appear in chains. If we found that they were targeted early on in their careers and then progressed through the ranks, this would cast doubt on the control argument. In contrast, under sponsorship, this is exactly what we would expect. People are sponsored at a time when they show promise but that promise has not yet been realized in high party ranks. Here we should find that the rank officials hold at the time they are involved in chains is lower than their highest rank.

This type of analysis requires us to shift from the person-level to the transfer-level, because only transfers are time-stamped and have time-varying party ranks. Table 1.6 presents results from hierarchical logistic regressions that predict whether a person is promoted into a higher party rank at the next Party Congress. Model 1 includes only our control variables (the year in which the transfer occurs, the age and party rank of the person at the time of the transfer, how many years are left until the next Party Congress, and the destination

of the transfer). Model 2 adds VC Length. Note that, according to a likelihood ratio test, model 2 performs significantly better than model 1, demonstrating the importance of our vacancy chain measure. The model suggests that a one-unit increase in the length of the vacancy chain increases the odds of being promoted at the next Party Congress by 40%. Thus, being part of a long chain captures a potential that is not yet actualized. It is a promise for the future. Note also that transferring to a provincial position is a good indicator of getting promoted in the future, as seen in the consistently negative effects for the other three organizational domains.[16]

The second analysis looks at effect heterogeneity across different types of people. We argue that under sponsorship we would see that younger elites benefit more from being involved in vacancy chains than older elites. To test this idea, we include an interaction effect between VC Length and Age (model 3, Table 1.6). The negative effect from this model means that the career boost that comes from being part of vacancy chains is higher for younger people than it is for older ones. Increasing the chain length by one unit increases the odds of being promoted by 11% for someone at the age of 60; it increases them by 146% for someone at the age of 40. This provides further evidence for the sponsorship argument.

This is confirmed by a closer look at the data. If we take all CCP politicians in our dataset who were involved in vacancy chains no later than 2005 while under the age of 45, the majority of them made it into the Central Committee. 32 out of 46 person-transfers had a maximum rank of Alternate Central Committee or above as of 2012. Some of the ones who did not make it into the Central Committee had career disruptions either due to death or anti-corruption efforts. In contrast, for party cadres who were involved in isolated transfers

16. One might argue that the effect we observe is merely the result of long chains containing important organizations and short chains containing less important organizations. The effect could then be the result of people in important organizations reaching high ranks. While this does not directly contradict the sponsorship argument, we here attempt to make transfers comparable by controlling for the type of the destination (ministry, province, party organization, or other), something that was not possible in the person-level regressions. Since this categorization of domains does not capture all meaningful differences between organizational units, we also provide results from an approach where we match transfers based on the exact same destination (Appendix A.1). The effect still holds.

Table 1.6: Transfer-level hierarchical logistic regressions predicting promotion in the next party congress

	model (1)	model (2)	model (3)
VC Length[a]		0.24***	0.30***
		(0.05)	(0.06)
VC Length[a] *** Age**[b]			−0.03***
			(0.01)
Age[b]	0.09***	0.08***	0.07***
	(0.01)	(0.01)	(0.01)
Year Rank	−0.38***	−0.41***	−0.35***
	(0.12)	(0.12)	(0.12)
Destination (= province)			
ministry	−0.51***	−0.45***	−0.43***
	(0.15)	(0.15)	(0.15)
party	−0.62***	−0.56***	−0.55***
	(0.21)	(0.21)	(0.21)
other	−0.84***	−0.72***	−0.69***
	(0.24)	(0.24)	(0.23)
Year	−0.00	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)
Years till Next Party Congress	−0.03	−0.03	−0.04
	(0.05)	(0.05)	(0.05)
Intercept	−0.61	21.10***	19.46***
	(5.35)	(5.73)	(5.57)
Transfers	2,492	2,492	2,492
Elites	1,482	1,482	1,482
Std. Dev. of Random Effect	1.14	1.1	1.08
Log Likelihood	−1,027.56	−1,017.27	−1,010.90
LR Test		20.57***	12.75***
Bayesian Inf. Crit.	2,125.50	2,112.75	2,107.83

Note: *p<0.1; **p<0.05; ***p<0.01

[a]VC Length is centered at 2; [b]Age is centered at the median.

under the same conditions (same period, same age cap), only 142 out of 551 person-transfers had a maximum rank of Alternate Central Committee or above as of 2012. Together, these results provide convincing evidence that a simple control strategy is not behind the transfers we observe. We begin to make sense of those patterns only when we think in terms of sponsorship.

CONCLUSION

Most work on political mobility in the Chinese party-state views mobility from the perspective of the individual. In this paper, we developed a complementary approach that views mobility from the perspective of the state. The core theoretical move is to think about status attainment not in terms of achievements by individuals but as a corollary of bureaucratic techniques. This switch encouraged us to consider the strategies pursued by the organizations that make hiring decisions rather than the attributes of those being hired. This is particularly important in cases where organizational strategies for personnel management change over time, as was the case in China after the reforms of 1978.

We used the concept of vacancy chains to study the structure of political mobility in China from such an organization-centered perspective. Vacancy chains turn the problem of where political elites aspire to work into a problem of where the organization needs them. We have argued that the patterns we observe are best understood as a residue of a strategy of sponsorship that identifies promising officials early on in their careers and places them on career tracks that ultimately lead to the highest offices in the party-state. Vacancy chains are not a magic exogenous treatment that boosts a party elite's career, but rather a signal, observed by us, that an official has been identified as promising and incorporated into the game of succession. It is as if vacancy chains are invisible bands that pull promising officials up the organizational hierarchy by connecting them to high-level elites at the top.

The organization-centered perspective encapsulates a different temporal vision of mobil-

ity. In the person-centered way of thinking, the past controls the present: a career move is a function of past experiences such as the acquisition of skills or credentials. In the sponsorship way of thinking, the future controls the present: a career move is a function of the promise the individual carries for the future of the party as a whole or factions within it.

The importance of sponsorship as an organizational strategy can be seen as a response to the demands faced by the modern Chinese state. In order to govern effectively, the Chinese state needs to do more than weaken powerful local elites. Keeping local officials at bay may have been a sufficient strategy in Imperial China. Empires, the structural analog to a holding company, are networks of power in which the center is only loosely connected to the periphery (Adams 1996; Barkey 1994, 2008; Go 2011). Under these conditions, maintaining stability and preventing local entrenchment might be enough. Modern centralized states, in contrast, face administrative challenges that require integration beyond simply weakening local elites. A pool of capable bureaucrats is necessary to penetrate society and carry out state policies (Evans 1995; Skocpol 1985). With an increasingly complex and vibrant market economy, the Chinese state no longer has a single objective as was the case during the planning era. Different regions design and experiment with developmental strategies that are best suited to their local needs. In addition, as the Chinese society becomes less restricted, the government encounters different societal needs, which are sometimes in conflict with one another. For instance, the tension between economic performance and environmental governance has been one of the most difficult challenges that Chinese local officials have to deal with in the reform era (Mertha 2009). To manage those increasingly complex needs, the state draws on capable officials who are not only good at performing single duties but also have the managerial skills to oversee complex tasks.

While in this paper we pitted sponsorship against a version of control that tries to keep officials from becoming too powerful, sponsorship can itself be understood as a form of social control. Weber (1958), H. Kaufman (1960), Doeringer and Piore (1979), Walder (1985,

1995a), and Iyer and Mani (2012) all point out that mobility serves as a means of controlling the members of an organization. By incorporating them early on, the CCP turns its elite members into organized men. It is this dynamic that Pareto (1970) had in mind when he suggested that a governing class might defend itself from individuals who could overthrow it by admitting them into the elite. By doing so, Pareto argued, these new recruits would adopt the attitudes and interests of the elite, thus turning enemies into allies. As such, our findings shed light on the authoritarian resilience of the Chinese party-state. As the enthusiasm for communist ideology faded, the party organization became the binding mechanism that regulated the behavior of Chinese political elites. And the movement of staff through the system of positions was a key element of those new dynamics.

CHAPTER 2

NEW WINE IN OLD BOTTLES: IDEOLOGICAL TRANSFORMATION AND THE RHETORICAL CREATION OF THE MARKET IN CHINA'S PEOPLE'S DAILY

Shilin Jia and Linzhuo Li

History is the process whereby the spirit discovers itself and its own concept.

— Hegel, *The Phenomenology of Spirit*

Chapter Abstract China's embrace of a market economy under the leadership of a communist party was one of the most surprising ideological transformations of the 20th Century. To reveal how the Party struggled to reconcile its internal ideological conflicts, we applied computational content analysis to the full text of more than 50 years of articles published in the official newspaper of the Communist Party of China, the *People's Daily*. Some remarkable internal patterns were discovered. During the course of China's economic reform, the Party's official state rhetoric progressed path-dependently at a highly consistent pace. The development and survival of new discourses crucially depended on the utilization of old repertoires. New words were always more likely to appear first in unusual contexts. However, without being absorbed internally, external political shocks could quickly die out in the system. In contrast, one mechanism of change always held; i.e., new elements were likely to survive only in existing stable contexts. Our results suggest that in cultural production, although novelty usually comes out of unusual combinations, adaptation in stable configurations (i.e., embedding new concepts in old ideas) is a key step in long-term cumulative change. In this light, market creation during China's economic reform can be seen as a product of a carefully balanced ideological project.

Introduction

In 1978, when the Chinese Communist Party (CCP) decided to initiate economic reforms across China, very few people could have imagined that communist China would one day become one of the world's largest engines of capital accumulation. After waves of political and cultural revolutions, the dominant ideology of China, defined by the CCP, went down a radically different path during the last two decades of the 20th Century. A party that once strove to adhere strictly to Marxist-Leninist orthodoxy gradually started to adopt and advocate economic policies that were very foreign to what it once believed. Throughout this period of reform, Soviet-style state planning was gradually discarded, state-owned properties were privatized on a massive scale, and the concept of a market economy was eventually enshrined in the Party's official ideology (Shirk 1993; Brandt and Rawski 2008). These reforms were followed by phenomenal economic growth and political stability among the CCP's ruling elites, which gave rise to China's appearance as a global superpower.

Most of the important work on the Chinese reform has focused on non-ideological dimensions of the reform. The trajectory of the reform can be explained in terms of changes of political power balance in CCP elite politics. The weak bureaucracy of the Chinese state after the Cultural Revolution gave Deng Xiaoping opportunities to maneuver pragmatic reforms (Shirk 1993; Vogel 2011). The decentralized political structure also gave space for regional experimentation (Qian and Weingast 1996, 1997; Oi 1999). Some scholars argue that once some market mechanisms are in place, the market would have an inertia to self-perpetuate and erode old institutions (Nee 1989; Nee and Matthews 1996). Broadly speaking, there is a common line of research that is in alignment with "the End of Ideology" thesis: market transition makes Communism and its ideology increasingly irrelevant. Since the end of the Cold War, the world has experienced a clear break from the past, and China has been integrated into a new global order (Fukuyama 1992).

Yet, how exactly could have a political party, with essentially the same body of people,

changed its tone and started to advocate for ideas that are in conflict with what they believed in the past? How could a new rhetorical regime supersede an old one? Did the party follow any internal logic in its change of tone?

If Communism and its ideology had indeed become increasingly irrelevant, then despite the radicalness of the rhetorical change, the change could have only been epiphenomenal. The party would have advocated for whatever beliefs that are consistent with its pragmatic interest, and there would be little internal logic to be found in its discourse change. However, any expert who is experienced with contemporary Chinese political discourses knows that the CCP has never loosened its hold on its communist legacy (Holbig 2013). Recent revival of Maoist tendencies in Chinese politics also suggests that Communist ideology is still pretty much relevant in today's Chinese politics and is perhaps never away (Yang 2014; Zhao 2016; Brown and Bērziņa-Čerenkova 2018; Klimeš and Marinelli 2018).

A minority group of scholars insist that there is an internalist story to be told about the Chinese reform (Tsou 1977, 1983; Sun 1995; Kluver 1996). They argue that the rhetorical transformation could have only unfolded step-by-step, and each step depended crucially on the conclusion of the last step. Without the gradual change in the definition of what socialism means, the market would have never become a legitimized concept. However, as human readers have a natural tendency to rationalize what they see, it is hard to say whether any internalist interpretation is a product of confirmation bias. Ideological and moral justifications could have mostly played a role in post-hoc sensemaking (Swidler 2001; see also Vaisey 2009).

More broadly speaking, the question at hand is whether ideology and culture have any independent explanatory power regarding social processes (J. Kaufman 2004). In sociology, this theoretical controversy has been best known in the debates about the cause of the Protestant Reformation. Max Weber famously argues that there exists "elective affinity" between ideas, and culture can be viewed as a semi-autonomous social sphere (Weber 1992,

1958). However, as the Weberian *verstehen* approach has gradually lost its methodological appeal, in the sociology of culture, contemporary studies have mostly adopted a social structuralist explanation of beliefs and attitudes. Even in cases where ideological and moral beliefs are not directly related to material interest, there is a commonly shared assumption that explanations do not lie in any intrinsic properties of ideas, and the associations of ideas can be best explained in terms of some extrinsic and coincidental properties such as the institutional environment (Meyer and Rowan 1977; DiMaggio and Powell 1983; DiMaggio 1997), the social status of their believers (Lamont 1992; Bourdieu 1993), or network homophily (DellaPosta, Shi, and Macy 2015). Some cultural sociologists insist that "cultural autonomy" exists but have been struggling with empirical validation (Alexander and Smith 2001). Establishing endogenous cultural associations with causal precision can be inherently a challenging task. Some internalist explanations rely on simulations that depend on the choice of modeling parameters (Goldberg and Stein 2018; Goldberg 2021). Some survey studies hinge on the possibility of internalist explanations of beliefs (Vaisey 2009; DiMaggio and Goldberg 2018; Amengual and Bartley 2022). However, survey responses usually lack details and resolutions and are not always best suited for generating insights about the inner logics of ideas.

Among internalist explanations of beliefs and ideas, one particular methodological and theoretical construct that has been proved very useful is the concept of a "discursive field." The discursive field can be thought of as a system of "possible ideas," that does not only delimit what ideas exist in a discursive space but also what range of ideas could have possibly been uttered or even entertained in a community (Foucault 1972; Wuthnow 1989; Spillman 1995; Snow 2004). In sociological analysis, it is especially useful for delineating the language space in which social changes are situated in (Wuthnow 1989; Bail 2012). And it adds a structural perspective to the theories of "issue framing" in showing how the framing of an issue can be more or less successful depending on the structure of meaning in which ideas are

embedded in (Steinberg 1999; Snow 2004; Fiss and Hirsch 2005; King 2007). With the help of commonly used dimension-reduction techniques, studies show that ideas usually do not span across all possible linguistic spaces but exhibit in clusters in a discursive field (Mohr 1998; Mohr and Neely 2009; Nelson 2021a). The emergence of large quantities of digitized texts and development of unsupervised word-embedding models have also made mapping more easily possible (Bail 2014; Nelson 2021b).

However, a precise operationalization of the concept is still lacking, and the concept mostly remains an abstract theoretical term. The "field" and the "space" although usually indistinguishable in their actual usage, should be two different concepts. The field is inherently a processual concept about how elements are prohibited, generated, or moved in space. It is inherently at odds with static approaches in representing elements of a space. Spatial clustering at best hinges on the possibility that a discursive field might exist but does not prove its existence. The internal logic of a system was still behind a curtain and remained elusive.

In this essay, we propose a dynamic framework to empirically validate the existence of a field in the rhetorical space of the Chinese Communist Party during its ideological transformation. We applied computational content analysis to more than 50 years of full-text articles published in the *People's Daily*, the official newspaper of the CCP. With the help of neural-probabilistic models, we constructed and aligned the word-embedding spaces of the words in the newspaper.

Our exploratory analyses reveal, first, that during the course of China's economic transformation, the Party's official state rhetoric evolved path-dependently at a remarkably consistent pace. Second, the development of new discourse and the survival of new lexicons crucially depended on the utilization of old repertoires. After the Cultural Revolution, the reform discourse first began by reverting back to an earlier base and developed gradually from that base. Later, controversial concepts, like "market economy", became stabilized only

by being contextually attached to an existing stable rhetorical neighborhood. The aligned vector spaces allow us to not only measure how stable each concept is in the space but also how stable the contextual space is at its different locations. It allows us to tell how the rhetorical system is moving and capture the presence of a field.

Then, following our manifesto that "a field only manifests itself in its effects," we show that the field is not only a useful analytical tool for describing what happened in the rhetorical space but also has a robust causal effect in determining which new concepts thrive and survive in the contextual space. Throughout the entire history of the *People's Daily*, field stability exerted a strong influence on the use of all words. Although new words were always more likely to appear in unstable fields, unless they were absorbed internally, external political shocks would quickly die out in this system. Only in existing stable fields, did new elements tend to survive. The effect demonstrates the existence of this field in which some ideas, but not others, are possibly spoken. Therefore, we argue that regardless of political turmoil, the Party's official language possessed an autonomy of its own.

China's Great Transformation in Retrospect

Marketization is a social process that institutionalizes the free flow of capital. As explained in Karl Polanyi's famous argument, the "free market" was realized in Western Europe precisely because of deliberate advocacy made by laissez-faire politicians and thinkers of the removal of any societal constraints on it (Polanyi 2001). Ideology is one type of such constraints that exists in people's minds. It can be broadly defined as a system of beliefs in which the beliefs are bound together to some degree (Converse 1964). It involves moral beliefs that regulate what should and shouldn't be done in society. Recent development in economic sociology points in the direction that the market does not only interact with morality, but as a performative concept, it is itself a moral object (Fourcade and Healy 2007). It demarcates what is exchangeable in society (Zelizer 1979, 1994) and is constituted

of institutions, tools, and calculative devices with built-in prescriptive theories (Callon 1998; Callon and Muniesa 2005; MacKenzie 2006). In communist China, it has always been a contested moral and political concept.

If someone opens an ideopolitics (*sixiang zhengzhi)* textbook used in Chinese high schools nowadays and finds relevant passages on the "socialist market economy," it would read something this: "China is still in an early stage of socialism, in which public ownership dominates but also co-exists with other types of ownership. Advancing the productive force is still the main task for Chinese society today. The market economy is a mechanism that regulates the distribution of resources to meet supply and demand relationships according to the Law of Value and is not exclusive to capitalism. It is an economic instrument that is suited for China's current level of productive force."[17]

To make Marxism and the market economy compatible, certain premises are clearly needed. The reconciliation relies upon an interpretation of Marx's historical materialism that puts a heavy emphasis on the determining role of productive force in the development of history. It requires the postulate that socialism has different stages, and there can exist some level of inequality and private ownership in some stages. It also needs to downplay some other Marxist principles, such as class struggle and anti-exploitation, both of which play pivotal roles in CCP's earlier revolutionary doctrines. How could have the CCP made the transition?

In order to appraise the importance of the question, some historical background is needed. The transformation of China in the late-20th century was a story of the transformation of the Communist Party. The CCP is an elitist party that was built upon the ideology of Marxism and Leninism. The Party is ideological and elitist by design. According to Lenin (1929), the working class does not possess class-consciousness on their own, and therefore the

17. The argument can be found, for instance, in Special Topic 5, Section 2 (pp. 83-86) of *Sixiang Zhengzhi Xuanxiu 2: Jingjixue Changshi* (Ideopolitics Elective 2: Basic Concepts of Economics) published by *Renmin Jiaoyu Chubanshe*. See also Ouyang (2001, 94–96) for an elaborate walk-through of the argument.

Communist Party should act as the revolutionary vanguard for the mass and lead the working class to the final victory of socialism. Within CCP party elites, ideological struggles about what path is the right "socialist" path to take are always a central theme (Tsou 1977). Mao Zedong famously argued that "class struggle should be emphasized every year, every month, and every day." During the Cultural Revolution, Mao sacked Deng Xiaoping three times because Mao was suspicious of Deng's pragmatist policy-approaches as "capitalist-leaning" (Vogel 2011; Tsou 1977). After Mao's death in 1976, Deng was able to take control of the party and initiate economic reform. However, Mao's legacy, as solidified in the so-called Mao Zedong Thought remained a central pillar in the Party's official doctrinaire while many party senior leaders such as Chen Yun still supported some adjusted versions of planned economy (Sun 1995). The reform was carried through a dual-track system with market-oriented mechanisms gradually replacing state planning (Shirk 1993; Naughton 1995). Its direction was not clearly laid out at the beginning and only became clearer after many rounds of in-party theoretical debates about the "essence of socialism" (Sun 1995).

In orthodox Leninism and Maoism, the "free market" has no place in a socialist economy. The CCP's later embrace of it is apparently a huge reversal in its official doctrinaire. However, existing studies of China's political economy have predominantly treated the rhetorical transformation as a consequence rather than a cause and focused mostly on non-ideological factors. Some scholars regard socialism in Post-Mao China only as an empty "disguise" of pragmatism (Cohen 1988; Pye 1986). Scholars holding the Market Transition Theory argue that the market unleashed itself through generating social institutions that eroded the old distributive system. During market transition, producers in the private sector gain a larger share of return relative to redistributors in state planning sectors, which will give the private sector incentives to continue to grow. The end result is a shift of resources and privileges from the public sector to the private sector. Therefore, the market has a force to self-perpetuate (Nee 1989). On the other hand, scholars studying elite politics point out the lack of a strong

state bureaucracy, an unintended consequence of the Cultural Revolution, as an enabler of change (Shirk 1993). The lack of centralization created an M-form of organizational structure. Fiscal decentralization and promotional incentives induced local governments to act as *de facto* enterprises in an industrial competition (Walder 1995b; Qian and Weingast 1996, 1997; Oi 1999; Zhou 2007). Some account also suggests that some earlier deviations from the planned economy started with local experimentation from below rather than directives from the top (Liu 1992). Ideology is at best absent, if not irrelevant, in these accounts. The legitimization of market concepts like "private ownership" and "property rights" is only a natural consequence of the institutional changes on the ground.

An implicit assumption behind many of these accounts is that party leaders can conveniently alter or abandon party discourses at will, and therefore ideology has no independent explanatory power. However, through careful historical analyses, some scholars have shown that the assumption is simply untrue (Tsou 1977, 1983; Sun 1995; Kluver 1996). In the formative stages of the Chinese economic reform, factional lines among the Party's top leaders were always drawn upon ideological lines, and policy debates were often justified in ideological principles (Tsou 1977; Sun 1995). Discourse changes were not post-facto justifications of policies already in place but results of vigorous debates. The rhetorical transformation took place through stages that depended on the conclusions of the previous stages.

For instance, the first major intra-party political struggle in the post-Mao era took place in the shape of an epistemological debate on the criterion of truth. After Mao died, Hua Guofeng, Mao's appointed heir, officially ended the Cultural Revolution but also proposed "Two Whatevers"[18] to consolidate his power base and maintain the status quo. To counter Hua's move, Deng Xiaoping and his allies (most notably, Hu Yaobang) endorsed and circu-

18. Hua made the statement that "we will resolutely uphold whatever policy decisions Chairman Mao made, and unswervingly follow whatever instructions Chairman Mao gave." The statement was printed in a joint editorial named "Study the Documents Well and Grasp the Key Link," in the *People's Daily* on Feb 27, 1977.

lated a philosophical article named "Practice is the Sole Criterion of Truth."[19] Referencing extensively the writings of Marx, Lenin, and Mao, the article advocated for using social practice as a fundamental and independent ground for judging and reevaluating communist theories. Despite the fact that Hua and his allies strongly opposed the article, the article triggered nationwide debates among party elites, which eventually turned into criticisms against "whaterism." Through this debate, Deng was able to garner enough support to form a large coalition and declared a decisive ideological victory in the Third Plenum of the Eleventh Party Congress held in 1978. The meeting and its conclusion marked the beginning of China's economic reform. Political calculations very likely played pivotal roles in the forming of the winning coalition, but the format of the political struggle was by any means ideological. (See Sun (1995, 22–34) for a full review of the debate.) The settlement of the debate also had immediate applications. The truth criterion allowed for a systematic reassessment of extreme Maoism. Through a subsequent debate, the party was able to drop "class struggle" out of the "principal contradiction of Chinese society" in the next year.

In the economic domain, one of the earliest controversial issues was how to deal with the communes. Although some villages in Anhui Province had already started delegating production responsibility back to households, the Party didn't change its official policy until 1981. In the early years, even reformers like Zhao Ziyang still saw household responsibility as a violation of the socialist economy (61). However, the earlier settlement on the Truth Debate gave space for local reformers to test what economic practices best meet China's "objective condition." Anhui's experiment proved successful in alleviating poverty, and eventually, the Party changed its official line and started to promote household responsibility nationwide.

In the next phase, the party line was divided between a conservative faction led by veteran leaders surrounding Chen Yun, who wanted to only make adjustments to the planning

19. The article was independently submitted by Hu Fuming, a philosophy professor at Nanjing University. It was first published in an internal party publication and was then reprinted in numerous party publications including the *Guangming Daily* and *People's Daily*.

system, and a radical reformist faction led by Hu Yaobang and Zhao Ziyang, who were in favor of pushing forward full-scale economic and political reform. Deng acted as the ultimate arbitrator in the middle. Throughout the 1980s, the reformers constantly pushed for reducing the role of planning while the adjusters were on defense. The two factions debated vehemently in the Party's official publications, such as the *People's Daily*, about the path that Chinese socialism should take.

The development unfolded in a series of theoretical debates about what constitutes the core elements of socialism. Major debates were centered around issues such as the legitimization of "individual economy," "self-interest," "commodity exchange," and "distribution according to labor." A recurring theme is reformers making the argument that the communal economy envisioned in Marx's writing is for society at a very advanced stage, and advancing the productive force of Chinese society according to the Law of Value should be the primary concern for Chinese socialism. On the other hand, the conservative faction constantly pushed back on the ground of economic overheating, bourgeoisie liberalization, and distribution of wealth. The end product is a zigzag process that oftentimes takes two steps forward and one step back (Shirk 1993). The political and ideological struggles ended with the 1989 Tiananmen crackdown, after which the conservative faction completely wiped their opposition out of CCP politics. However, the earlier debates weren't futile. Three years later, when Deng intervened and re-activated the market reform, many of the concepts and elements were already there for a "socialist market economy" to emerge.

Scholars who gave such internalist explanations rightly pointed out the importance of ideological debates in this historical transformation. However, a skeptical reader could always challenge the veracity of an ideological interpretation. To social scientists who insist on giving objective explanations of the world, the term "ideology" has an inherently deceptive connotation (Geertz 1973, 193). It is also hard to entertain what the counterfactual would be. Although many of the important articles during the period, including the Truth article,

cited Marx, Lenin, and Mao extensively to justify their positions as truly socialist, just based on qualitative readings of the articles, it is hard to surmise whether the history would have been any different had the arguments been made differently.

The issue here corresponds to a century-long problem in the sociology of culture, which is best known in the debate surrounding the causal role of the Protest Reformation. In many respects, China's redefinition of socialism is comparable to the Reformation of the 16th century. Both movements were a departure from a society in which a set of ideological principles provided guidance for every aspect of its social and economic life. Both had an institutionalized clergy that adopted a highly confined vocabulary. From their own vocabs, the reformers were nevertheless able to develop some new discourses that fundamentally changed what was considered good vs. bad in their existing value systems and reoriented people's social and economic life. However, regardless of their significance in the history of ideas, what causal roles they played in the development of history is an open question. In the next section, we will go back to the theoretical debate in sociology and try to tackle the root of the problem.

Protestant Reformation Revisited

From a Marxist perspective, an ideological change takes the form of disillusion and has no significance of its own. The historical development of a society is determined by the material relations in its economic base. The ruling class imposes its ideology as the ideal of universal mankind. Ideology is a "false consciousness" that mystifies the dominant relationships of production. Any change in ideology only reflects a change in the ruling class. As Marx (1974) states in *The German Ideology*, "morality, religion, metaphysics, all the rest of ideology and their corresponding forms of consciousness no longer retain the semblance of independence; they have no history and no development."

Since the founding of sociology, sociologists have always been deeply concerned about

whether there is anything beyond the materialist thesis, and the Reformation has been repeatedly used as a case to generate and test theories of social and ideological change. In *The Protestant Ethic and the Spirit of Capitalism*, Max Weber (1992) famously argues that historically there existed a causal relationship between the Calvinist idea of "predestination" and the early development of Capitalism. According to Weber's thesis, Calvinist believers developed a work ethic of wealth accumulation due to their desire to demonstrate that they were elected by God. Calvinism was conducive to the early development of Capitalism because there was an "elective affinity" between Protestantism and Capitalism.

The historical accuracy of Weber's account remains to be a subject of debate. Weber might have cherry-picked some of his evidences (Walzer 1966). Given that Capitalism was also developed in other parts of the world where the Protestant Ethic was absent, it remains doubtful whether the role of the Protestant Ethic was truly decisive in the development of Capitalism (Swidler 1986). However, on top of this first causal layer, Weber also poses another type of causal question, which is concerned about the inherent connectedness of ideas in ideological transformation. This second vision is more clearly articulated in his piece "Religious Rejections of the World and Their Directions" (Weber 1958). It has less to do with whether the Protestant ideology caused any economic change. The question is whether ideology would change in predictable directions when it is in conflict with other spheres of activities. Weber sees the Protestant Reformation as a solution for reconciling the conflict between the universal brotherliness of religion and the this-worldliness of economic calculation. Weber argues that although humans are not necessarily uncomfortable with holding inconsistent beliefs, the establishment of priesthood and separation of social activities would nevertheless put each social sphere at a tendency of rationalizing itself. And a serious problem would arise when the rationale of one sphere contradicts that of another sphere. In such a situation, the problem demands a solution. Because rationalization imposes constraints, there is only a limited number of directions in which change could take place. Weber argues

in precision that Puritan asceticism was one of the only two ways of escaping the tension, and therefore the ideological change was predictable and followed an "inner law" of its own.

Ann Swidler (1986) challenges the idea that culture guides people's actions in predetermined directions and questions Weber's ability in explaining why the Protestant Reformation proceeded in the ascetic direction rather than other possible directions. Swidler also cites the study of Michael Walzer (1966), which puts the emergence of Calvinism into its historical context. She points out that Calvin was not just a revolutionary priest but also the ruler of the small theocracy of Geneva. His doctrines served his need of maintaining organized control of his disciple citizens. Swidler proposes an alternative theory of seeing culture as toolkits for people to organize their actions to meet their situational needs. Although Swidler generally dismisses the internal coherence of culture, in the same article, Swidler also makes a distinction between "settled lives" and "unsettled lives," and she at least suggests that during unsettled periods when great social transformations are ongoing, different ideologies emerge in competition with each other and articulate in more self-consistent fashions.

A closer read of Walzer (1966) nevertheless suggests that there was more beyond a structural need. Walzer dismisses the Weberian argument that there was a historical link between Calvinist theology and the idea of capital accumulation as there was very little historical evidence suggesting that religious reformers at the time were preoccupied with economic thinking. However, the Calvinist saints were also not just radicalized revolutionaries. They belonged to a rising social class who had an existential need to secure a spiritual safe harbor in a world of disorder. The need was not just materialistically structural but also psychologically structural. Through the re-discovery of "predestination" in the lines of the Bible, they found a way of orienting their spiritual life in a transformed world. In this sense, "the saints discovered in themselves a predestination." Walzer's account suggests that the coherence of Calvinism was not a byproduct; rather this ideological movement had an inner core of its own. In this sense, although he rejected Weber's first causal layer, his account is still

consistent with Weber's second layer.

Robert Wuthnow (1989) tries to make a synthesis by theorizing the tension between the material and psychological as a problem of articulation. He reasons that in order to make an ideology achieve a long-lasting influence, a reformer needs to both articulate his ideological vision well to serve certain social structural needs and also transcend from the historical specifics to make the ideas universal. Wuthnow gives a detailed analysis of the social context in which the Reformation arose. In short, Europe's commercial expansion over the sea gave rise to a new type of political entity, the cities. The reformation was a coalition between the city princes and the religious reformers against the Catholic moral order of the landlords. Martin Luther's doctrine of individualism served the need for breaking religious believers' attachment to the landed nobility. However, the social-structuralist account was supposedly only half of the story. As the religious reformers started to enjoy stable access to resources, they also enjoyed an autonomy of their own. They provided guidance for specific actions only in accordance with a general symbolic framework defined by a set of basic binaries such as "good vs. evil," "temporal vs. spiritual," and "man vs. God." Wuthnow calls such a framework a "discursive field" and points out that Luther and Calvin each had a distinctive field that was guided by its internal structure. As the movement unfolded, the fields started to acquire some independence from the social environment. However, despite his remarkable theoretical effort, Wuthnow's book provides much fewer empirical justifications for the second half of the story than the first half.

The literature on the Protestant Reformation suggests that there are two important questions that can be asked about the role of ideology in social transformation. The first one is concerned with the causal link between the ideational and the social structural. The second one is about the internal dynamics of ideas. The question is whether ideological change can be explained and predicted from within. Neither question is settled. But as exemplified in studies of the Protestant Reformation, accounts given to the second question

are oftentimes much harder to articulate. In this essay, we are primarily concerned with the second question.

Internalist Explanations of Culture

Among cultural sociologists, although there exists a wide range of theories with regard to how culture interacts with the materialist world, it is, in general, less controversial to admit that there are at least some internal structures in cultural systems. For instance, in "Culture in Action," although the main argument is that culture, rather than providing end-values, is "used" as "toolkits" in people's everyday life, Swidler is very careful in distinguishing her theory from a reductionist account and puts a special emphasis on how culture as repertoires can be constraining. The controversial matter is to what extent do internal structures exist and matter in social change. In Swidler (2001)'s words, the problem of "whether and how some cultural elements control, anchor, or organize others" remains "the biggest unanswered question in the sociology of culture."

Rooted in the works of Weber (1958) and Durkheim (2001), sociology has a tradition of explaining how cultural systems operate from within. However, as the discipline becomes further established as an empirical science grounded in objective rigor, it also faces strong methodological and theoretical resistance toward treating meaning too seriously (Geertz 1973, 195). For instance, although social contexts and discursive fields allegedly play equal weights in Wuthnow (1989)'s theory of the Protestant Reformation, Wuthnow devoted far more work to the former rather than the latter. One reason is perhaps that the former is more easily empirically verifiable than the latter. In his methodological note, Wuthnow criticized classical theories for treating culture and ideology too vaguely as wholistic world-views (or "free-floating Zeitgeist") that have no "empirical referents" and dismissed Weberian psychological explanations as subjective (517-526). In *Meaning and Moral Order*, Wuthnow (1987) advocated for abandoning the problem of meaning altogether in cultural analysis.

According to J. Kaufman (2004)'s review, recent sociological developments in internalist explanations of culture have taken a post-hermeneutic orientation and departed away from the interpretation of meaning. One notable attempt is the semiotic approach led by Alexander and Smith (2001). As strong proponents of "cultural autonomy," Alexander and Smith are only against "thin" or "weak" hermeneutic analysis that reads a text at its face value. Following Geertz (1973), they proposed to read social texts as coded symbols rather than reified abstract values. Unlike Geertz, who refuses any kind of generalization, Alexander and Smith propose to follow the structural linguistic tradition of Saussure (2011) and Lévi-Strauss (1983) to study culture as a system of "arbitrary" signs, which are not necessarily derived from any social referent, or the signified, but only from their own relationships with each other. They argue that structural analysis has great analytic power to yield generalizable causal claims. For instance, Alexander and Smith (1993) find that there are a basic set of "democrat codes" in American political discourses, and this set of symbols get reproduced over and over again in contingent political contexts over the centuries. In another study, Alexander (2002) delineated how the word "Holocaust" detached from its initial usage and became a free-floating signifier in American culture. His structuralist explanation is that the word only achieves its prominence by becoming an anchor in a symbolic system so that it can be used for making symbolic analogies with other words. Alexander (2002) agree with Kane (1991) that the analytic autonomy of culture is only situated in concrete historical contexts. However, because culture forms an analytically distinct social domain, they argue that the causal autonomy of culture can be studied separately from the comprehensive problem of how culture interacts with other social structures (196). They nevertheless admit that their qualitative approach does not allow for rigorous counterfactual falsification (167).

Another approach that has achieved a relative degree of causal success is the ecological analysis led by Lieberson (2000). Lieberson sees cultural symbols are in competition with each other. When one symbol is overused, novelty-seeking consumers would start to look for

alternatives. In his empirical analysis of American baby names, He found that the naming practices followed certain internal fashion cycles. Because new tastes were only built on existing old tastes, changes only took place in incremental fashions. The trends can be predicted in empirical observations. However, this embryonic form of ecological analysis does not account for any web of meaning cultural symbols are embedded in.

In general, internalist explanations of culture hypothesize a relative degree of autonomy in cultural production. Despite the fact that culture is constantly influenced by external social factors, an internalist account only needs to demonstrate that there are some internal causal mechanisms in the production of cultural symbols. Furthermore, a causal claim can only be rigorously made if its hypothesis is falsifiable, and its outcomes are predictable. Following Alexander and Smith and Lieberson, our empirical framework employs both 1) a structuralist way of analyzing ideas as symbols (words) and 2) a falsifiable ecological analysis. And we aim to answer one simple yet fundamental question: what internal structure of an ideological system constrains and enables certain ideas to thrive or perish? In the next section, we will put a special emphasis on the concept of the field and explain how it can be a useful analytic device for studying our problem.

Discovering Self-Organization in Fields

In sociology, the concept of "field" has its most famous application in Pierre Bourdieu's analysis of cultural production. Bourdieu (1993) sees cultural fields as arenas of competition for social actors to reproduce their social distinction. Every actor has a position in the arena relative to others, and the field is the topological space in which the positions are in. The field has an effect in the sense that every position induces a disposition (habitus) that gives the person who occupies it a sense of the world he is in. Although Bourdieu is always more interested in the production of culture rather than internal structures of culture, his conceptualization of "field" nevertheless provides a powerful analytical tool for making

endogenous explanations of social phenomena as he argues that fields are objective realities and induce rules of their own (Bourdieu and Wacquant 1992, 94–114). The concept has also been widely used in the studies of organizations.

John Levi Martin (2003) proposes that field theory could have a wide range of applications in social sciences. As the concept was first rigorously applied in the studies of classical physics, it is useful for explaining phenomena in which elements 1) are not in direct contact with each other and 2) experience forces induced by their positions. A field or field effect exists insofar as it offers a parsimonious explanation of the movements in it. Just like in the case of an electromagnetic field or gravitational field, a field is not a separate entity by itself and is only construed by the things in it. But it can nevertheless be self-organizing, and its self-organization can be demonstrated from the patterned effects observed in it. Field can be a useful analytical device for studying meso-level phenomena when the micro-level actions cannot be exactly specified. The lack of micro-level links is not necessarily a defect as it allows researchers to analyze problems from a holistic perspective. Martin argues that, in sociology, a field approach can be especially fruitful when it is applied to some existing line of research, such as the "inner laws" of social spheres specified by Weber. For all these reasons, field theory is a powerful analytic tool for studying endogenous social phenomena.

Just like in organizational studies, the unit of analysis in a social field does not have to be individuals but can be any meso-level social construct. In sociology, field as a concept has already been deployed in internalist explanations of culture (J. Kaufman 2004). For instance, in Lieberson (2000)'s ecological analysis, the fashion cycles of baby names can be thought of as a one-dimensional field phenomenon. In *Chaos of Discipline*, A. D. Abbott (2001) argues that there is a "fractal process" in the internal division of academic disciplines. The fractal process can be thought of as a field process in a discrete space. Wuthnow (1989) deploys the concept of "discursive field" in his explanation of the Protestant Reformation. In the social movement literature, "discursive fields" have also been used for a better understanding of the

framing processes of social movement discourses (Snow 2004). Ideas and cultural symbols are meso-level social constructs that do not directly interact with each other. Their usages are always mediated by people. Nevertheless, they exist always in relationship with each other and form an analytically distinct domain. They also occupy ecological "positions" relative to each other. Their usages can be affected by their internal "field" positions.

Despite its usefulness, the concept of "field" has mostly been used metaphorically, and its operationalization remains elusive. In this essay, we treat words as signs in a symbolic system and propose a research design to study the movement of words in their contextual field. One less clarified problem in the literature on field analysis is that fields are oftentimes operationalized as static constructs although field theory is intrinsically dynamic. Structure and change are paired concepts. Social systems are never in homeostasis. To study structure is to study how change is enabled and disabled. And the empirical question we want to focus on is: where do changes survive in a contextual field?

Another unclear problem in the literature of field studies is how boundaries should be defined (Martin 2003, 24). We take it to be less of an issue. The existence of a field only manifests in its effects. We do not assume that a field constitutes a separate reality. In this study, nor do we need to make a Parsonian a priori assumption about the analytic autonomy of culture. We only argue that if cultural autonomy is present, it should be observable in field effects.

We put a heavier emphasis on the survival rather than the introduction of changes. We do not hold that ideology is a closed system. As ideology is always influenced by non-ideological forces, changes are constantly infused from the outside. The introduction of change is a non-issue. Cultural evolution happens in the long run, and what matters is what survives. Social fields thus differ from physical fields in one crucial respect: social fields are ecological. In the social world, fields are arenas of contestation in which elements struggle for survival. In the short run, changes could come from random disturbances and external shocks. However, in

the long run, if a field position induces an effect, the effect should manifest in the survival of the elements that occupy the position.

In particular, we propose that a new word has a better chance of survival if it can be incorporated into a coherent rhetoric field. Social phenomena are organic and living. Living organisms reverse the law of entropy and create orders and structures in the world. In the production of social texts, self-organization could emerge as a consequence of rationalization or structural differentiation of signs. Rationalization is a process in which ideas become internally coherent with each other (Weber 1958). Coherent ideas are more powerful as they are more persuasive and universal. There is no reason to assume that people always think rationally and coherently. But as ideas compete with each other, there is good reason to believe that, coherent ideas should hold competitive advantages and have better chances to survive (Swidler 1986). Meanwhile, ideas are also expressed as symbols. In a symbolic system, signs only become symbols as they are used in structural differentiation with each other. A sign (word) wouldn't have any symbolic value if it is randomly used in combination with others. To achieve clarity, signs also need to be coherently used. In field terminology, the coherence of a rhetorical field can be represented by the stability of the relationships of the elements in it, i.e., the stability of the field. To test whether cultural autonomy is present, we hypothesize that in ideological change, - New words are more likely to survive in stable fields.

In the next section, we will introduce our research design and explain how we operationalize the construction of a field and test its effect.

Research Design

Data source

The discourses we studied are more than 50 years of full-text articles published in the *People's Daily*. The *People's Daily* is the ideal source for studying the CCP's ideology. It is the official organ of the Central Committee of the Communist Party and publishes the most authoritative statements regarding the Party's political and economic policies. Even top Party leaders (including Mao Zedong and Deng Xiaoping) have occasionally participated in writing and editing articles (Wu 1994). Rather than acting as a media "watchdog," it is a de-facto "propaganda machine" (Brady 2009). Having first been published in May 1946 and having a current circulation of 3 million, the *People's Daily* is one of China's most influential media outlets, and UNESCO lists it among the world's top-10 news media. Its influence often extends to other media and organizations in China. For instance, the paper's commentaries often become the main material for nationwide weekly "political studies," and other media outlets, including radio, television, and newspapers, often have to re-broadcast or reprint important commentaries that appear in the *People's Daily* (Wu 1994). In Mao's era, editorials and leading articles can also play a major role in launching new political campaigns (Brady 2009).

All of its issues can be accessed directly on its official website. Our data came from the database of *wengewang*, a third-party website that purchased the full-text from the very first issue published in 1946 to the last issue of 2003. We checked the completeness of the database by matching all the article titles in it with those on the official website. The overlap rate was at least 99.4%. The entire corpus contained a total of 1,342,414 news articles and commentaries, encompassing 260 million semantic words and a set of 1,214,510 unique vocabularies.

Exploratory method

A direct way to detect signals in comparisons of texts is to count word frequencies. In computational content analysis, the analysis of word frequencies has been widely applied to domain-specific keywords to track substantive changes and differences in texts (Gentzkow and Shapiro 2010; Klingenstein, Hitchcock, and DeDeo 2014). It can also be applied to meaning-free functional words to track changes and differences in linguistic and literary styles (Danescu-Niculescu-Mizil et al. 2012; Hughes et al. 2012). In this study, we are more interested in changes in substantive content rather than linguistic styles; thus, only non-functional words are used in our analyses (See Supplementary Section 1). After the frequencies of all the words of interest are counted, a probability distribution of the words is derived as a bag-of-words representation of the corpus. Temporal changes are our primary focus. The probability distributions of words between two time periods are compared by calculating the Kullback-Leibler (KL) divergence. Furthermore, we employ correspondence analysis (CA) (See Supplementary Section 2) to term frequency-inverse document frequency (tf-idf) matrices (See Supplementary Section 2) constructed from the word frequency distributions. CA is a dimension-reduction technique that is commonly applied to count data. It allows us to capture the most significant trends in changes in word frequency.

Construction of field

Field analysis presupposes the existence of a social space (Martin 2003). To capture nuanced semantic relationships that are ignored in our bag-of-words models, we employ a neural probabilistic model to embed words into high-dimensional vector spaces. In neural probabilistic language models, words are learned as vectors in a hidden layer of neurons to predict words that appear around them in a corpus (Bengio et al. 2003). This class of methods has achieved tremendous success in natural language processing. In this study, we used the word2vec model invented by Mikolov et al. to train our word embedding.

Word2vec is a very fast and efficient algorithm for learning high-quality word representations (Mikolov, Sutskever, et al. 2013; Mikolov, Chen, et al. 2013), and it has become a state-of-the-art technique in computational content analysis (Caliskan, Bryson, and Narayanan 2017; Garg et al. 2018; Kozlowski, Taddy, and Evans 2019). In general, the model is based on the distributional hypothesis in linguistics, which asserts that words appearing in similar contexts hold similar meaning (Harris 1954; Firth 1957; Mikolov, Sutskever, et al. 2013). It resonates well with structuralist theories in structural linguistics and cultural anthropology (Saussure 2011; Lévi-Strauss 1983; Kozlowski, Taddy, and Evans 2019). Its success suggests that a machine is able to produce natural languages by just learning how words are used with other words.

This structuralist way of thinking is also fundamental in semantic network analysis (Rule, Cointet, and Bearman 2015). However, networks rest on discrete spaces, and the dimensionality of a semantic network grows linearly with vocabulary size, which can make it very difficult to analyze. Word embedding provides an optimal way for learning semantic geometries in a moderate number of dimensions with a low informational loss. It also allows one to perform algebraic operations on its learned vector space to reveal semantic compositionality and analogy (Mikolov, Sutskever, et al. 2013; Kozlowski, Taddy, and Evans 2019). Supplementary Section 3, contains some well-illustrated examples of vector algebra generated from our models. Vector fields can thus be easily modeled on top of word-embedding spaces

Another benefit of word-embedding spaces is that they provide a way for us to select domain-specific keywords. More specifically, we use the nearest neighbors of the words "economy" and "politics", respectively, as economic keywords and political keywords. These domain-specific words allow us to focus on specific substantive areas in our exploratory analyses.

Word embedding spaces only preserve the similarity between words and are, therefore, coordinate-free. Nevertheless, vector spaces in different time periods can be aligned to track

rhetorical changes over time (Kulkarni et al. 2014; Hamilton, Leskovec, and Jurafsky 2016b). We used Procrustes Analyses to align the vector spaces of all the years to one coordinate system to permit comparisons. This allowed us to find the (in)stability of a word w from year $t-1$ to t as the cosine dissimilarity between the vector representations of the word in $t-1$ and t. Moreover, because words that are close to each other in a vector space are words that share similar contexts, the method allowed us to measure not only the instability of every word but also the instability of the contextual field the word was in.

To calculate the field instability of any position in the vector space, we tale its k-nearest neighboring words in the vector space in year t. Then, we compute the instabilities of the neighboring words from $t-1$ to t and take their average as the instability of the field position. Alternatively, field instability can also be measured as the substitution rate of a position's number of new neighbors over its number of old neighbors in year t in comparison with $t-1$ (Palla, Barabási, and Vicsek 2007). A field is stable if its neighbors are relatively the same over years. This alternative measure does not rely on Procrustes Analysis and can be used to check the robustness of our analyses.

Figure 2.1 gives an example of a stable field and an unstable field under each measure. Stable fields are neighborhoods where there are fewer movements and fewer new neighbors. Conversely, unstable fields are neighborhoods with more movements and more new neighbors. Our measures should make an intuitive sense.

Exploratory Findings

Path Dependency with Critical Junctures

Path dependency has been found to be a central feature in cultural evolution (Hughes et al. 2012). Novel rhetoric cannot be created out of nowhere and depends heavily on time trends and what is available from the past. We found persistent path dependency in the

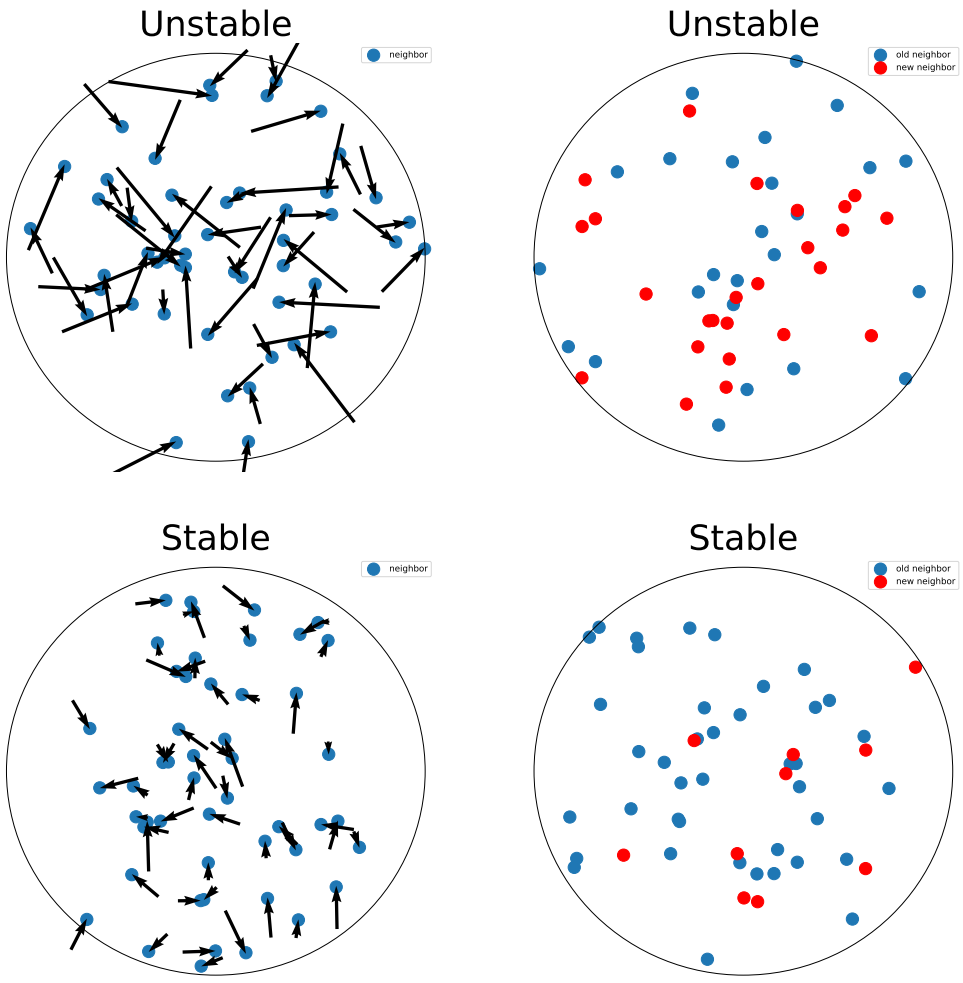

Figure 2.1: 2D examples of stable and unstable fields under two measures of field stability

CCP's state ideology, especially in the economic domain. Heatmaps, based on year-to-year KL divergences, are shown in Figure 2.2. The analysis is applied to all words as well as domain-specific words. Three major phases of the People's Republic of China can be clearly identified in all three heatmaps. The three phases are the socialist construction period from 1949 to 1966, the Cultural Revolution period from 1967 to 1976, and the economic reform period from 1977 to the last year included in our analysis. It can be seen that the year-to-year KL divergences in the last period were especially small in comparison with the earlier two periods. Although this period was the time when the great economic transformation took place, it was a very self-consistent period in terms of party ideology.

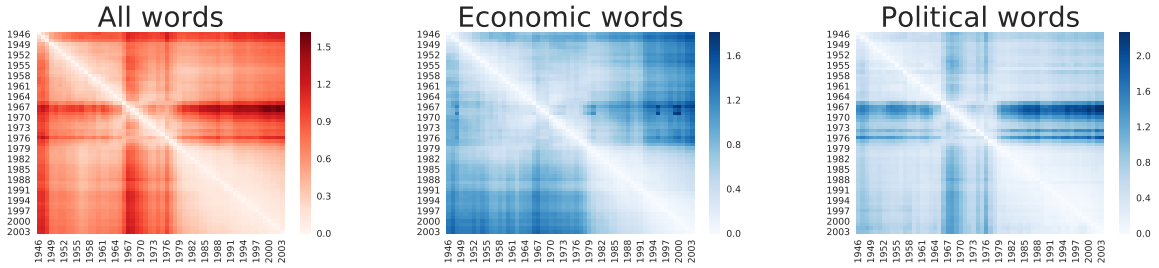

Figure 2.2: Heatmaps based on year-to-year KL divergences

Time series, which are constructed with two choices of prior distribution, are plotted in Figure 2.3. The time series on the top of the figure are constructed with the prior distribution equal to the probability distribution of words in the final year (i.e., 2003). The time series at the bottom are constructed with the prior distribution equal to the probability distribution of words over the 12 months preceding month t. As shown in the graphs, after the Cultural Revolution, the CCP's official state rhetoric had, in general, moved in a very smooth and linear fashion across almost the entire time. This time series has a much shorter memory during the reform period than the previous period. After an external shock, it appears to quickly revert back to its mean. For instance, in Figure 2.3, an obvious external shock can be detected at t =1989M06, which precisely corresponds to the crackdown on the Tiananmen Student Movement. The word frequencies in that month differ radically from the frequency

distributions in the preceding months. However, unlike political events during the Cultural Revolution, the Tiananmen event has very little long-term impact on the CCP's overall state discourse. In the economic domain, the shock is hardly detectable. Quickly after the crackdown, the CCP's official state rhetoric proceeded as if Tiananmen never happened.

In contrast to the overall stability, patterns in the economic domain are somewhat different. There is no obvious shock during the entire reform period in terms of monthly divergence. However, in terms of convergence toward the present, it can be clearly seen that the period can be further divided into two sub-regimes. The critical time point is the beginning of 1992 and corresponds to another critical event in the history of the CCP, a southern tour made by Deng Xiaoping, the chief architect of China's reforms.

Time dependencies can also be clearly captured by CA. Two CA biplots are shown in Figure 2.4. One is applied to all economic words at all time points, and the other is applied only to economic words in the period from 1982 to 2003. CA is completely unsupervised. Two words or two months would appear close to each other only if their respective columns or rows are similarly loaded. Arrows are drawn to connect the centroids of adjacent years. In the first plot, time dependencies could be traced with two dimensions. In the second plot, only the first dimension is needed to align all the years in a linear order. Unlike in previous years, the CCP's economic discourse never went backward in the last two decades covered by our analysis.

However, with the help of the second dimension, a change in discourse could again be detected. In the spring of 1992, Deng Xiaoping, who was officially retired but still the most powerful political figure, made a series of public visits to several special economic zones in South China and strongly signaled his dissatisfaction with the progress of China's reforms. Authorities in Beijing eventually responded positively in the *People's Daily*. The event reactivated China's market reforms (Vogel 2011). Although the change in tone was pushed from the outside, our result suggests that it was not so much an external shock as it was a

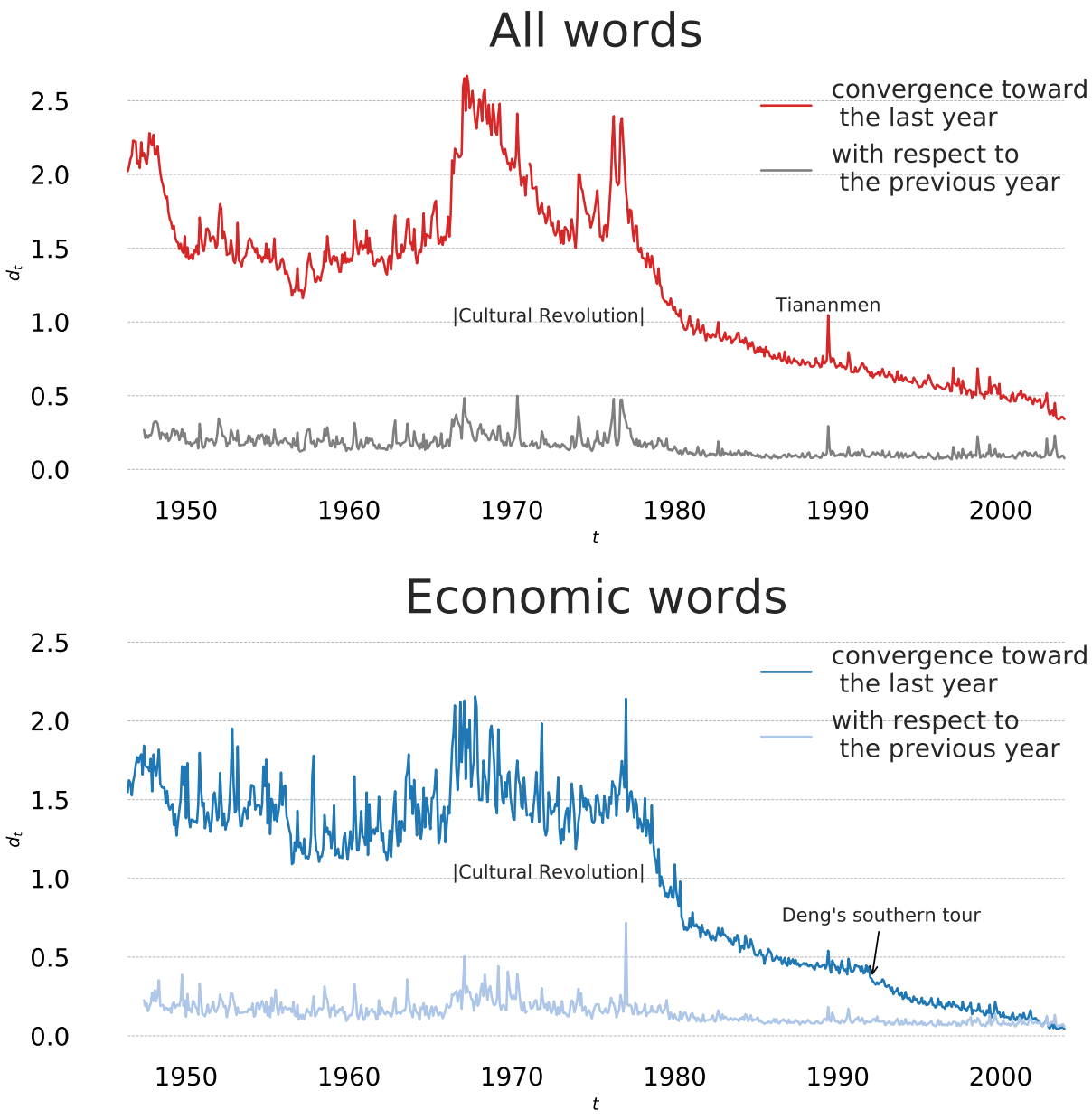

Figure 2.3: Monthly frequency change measured in KL divergence

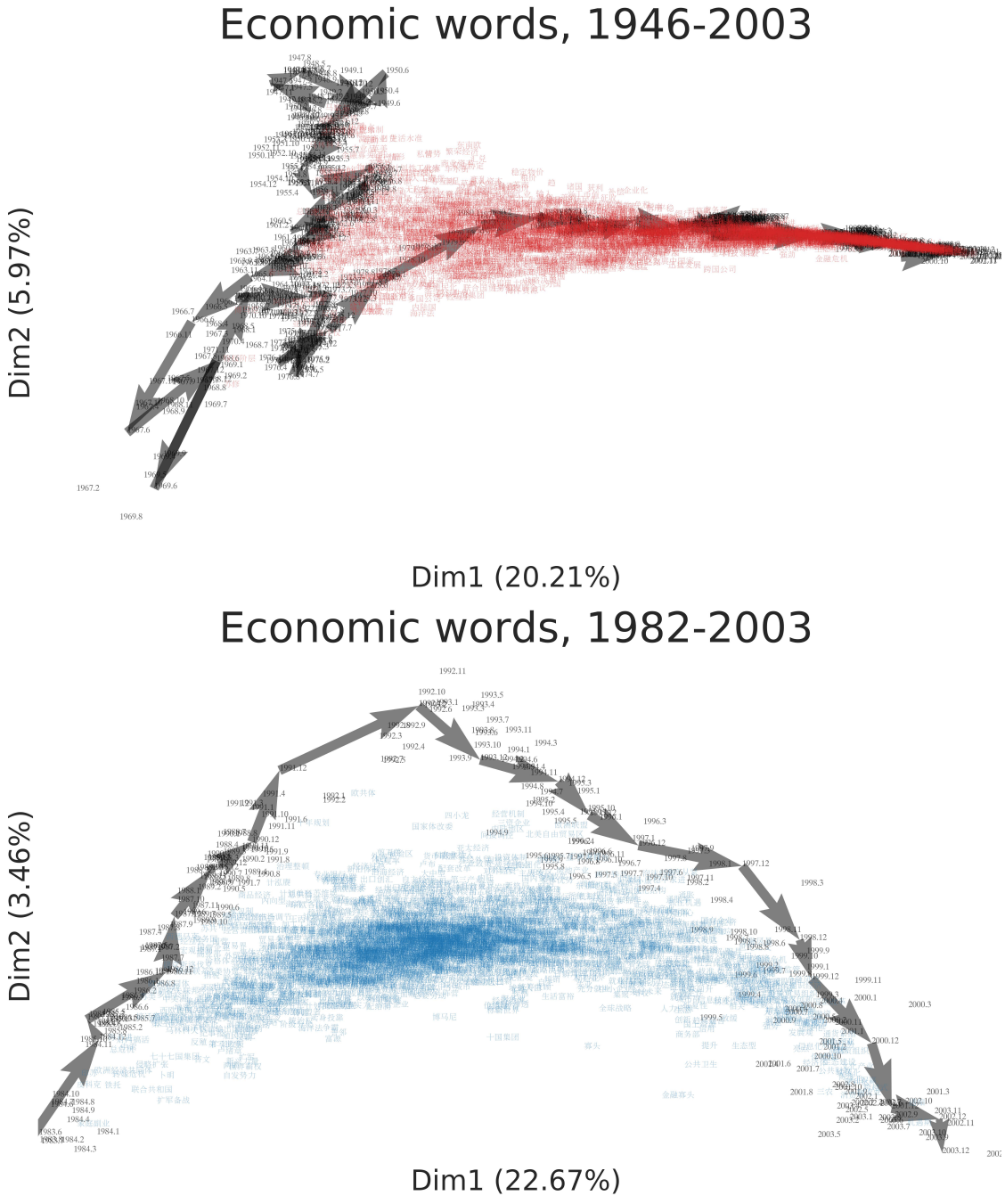

Figure 2.4: CA biplots in row principal components, dim 1 vs. dim 2. In the graph on the left, the first principal dimension captures an ideological shift from the leftist point during the Cultural Revolution toward the latest time point in the reform period. The second dimension captures how the Party's state economic ideology became radicalized during the Cultural Revolution and then reverted back to its 1950s level before embarking on its journey toward reform and opening-up. A biplot of economic words from 1982 to 2003 is shown in the right graph.

trigger of change in internal patterns. More about this will be explained in the next section.

New Wine in Old Bottles

Next, we find that the entire transformation was initiated in the late 1970s by initially utilizing some existing 1950s repertoires. The CA biplot shows that the CCP's official state ideology after extreme radicalization during the Cultural Revolution, first, shifted gradually back to the level of 1950s (in a recovering period) in both the first and second dimensions, before evolving to its current form.

We also measure yearly changes in discourse in the word-embedding spaces. First, we select the words most similar to "economy" and "politics" in the last year. Then, we extrapolate the probability of their co-occurrence in all previous years. The results, which are shown in Figure 2.5, are similar to what we find in the word-frequency analysis. In both the economic and political domains, there is a "recovering" process. Some scholars argue that the Cultural Revolution destroyed China's state bureaucratic machine and unintentionally paved the way for later institutional reforms (Shirk 1993; Naughton 1995). Our results suggest that, rhetorically, the Party first changed its discourse back to its earlier base and proceeded from there. The results are consistent with some earlier interpretive studies of the CCP's ideological change (Tsou 1983; Sun 1995).

Then, we zoom into what happened in 1992. We find that when a sudden change happened, newly introduced concepts still needed to be attached to existing stable repertoires to become stabilized. Figure 2.6 is a visualization of the contextual field around "market economy" in 1992. The blue arrows represent the movement of the phrase in the contextual space. The movement suggests that since its inception in 1984, the phrase was used in different contexts until 1992. After 1992, its contextual usage stabilized, and the phrase finally settled down in one location. The figure also suggests that when the phrase was first introduced, it was used in unstable fields (as shown along the z-axis) and surrounded

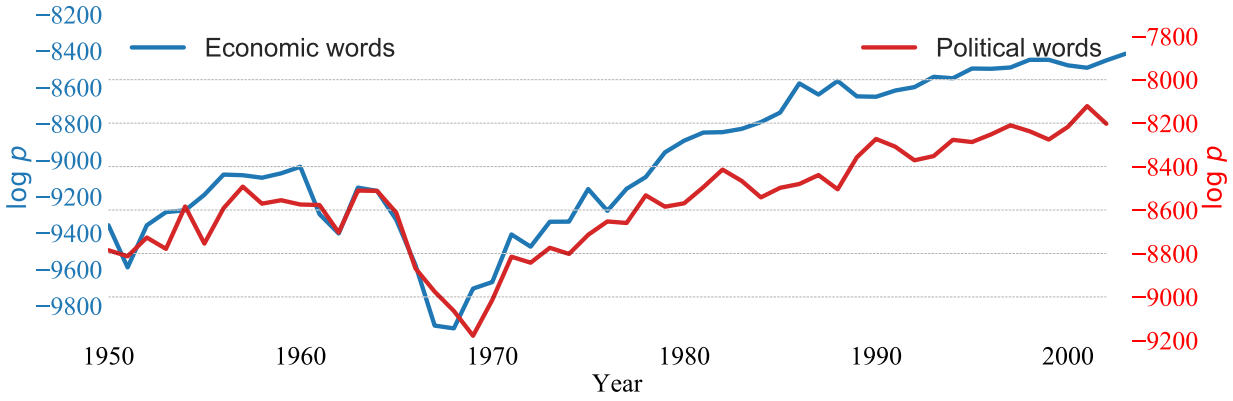

Figure 2.5: Log-likelihoods of co-occurrences of the most recent year's political and economic words given each year's corpus.

by words that were fairly sensitive words, such as "monopoly", "multi-party system", and "financial crisis". The phrase became stabilized in 1992, by being contextually attached to a highly stable subspace of words, surrounded by safe words like "reform" and "socialism". Figure 2.7 corroborates the fact that its new contextual neighbors were already very stable prior to the year of stabilization, compared to its old neighbors.

Confirmatory Analysis

The settlement of "market economy" in 1992 is consistent with our hypothesis. The concept only thrived in the rhetorical field after it was incorporated into a stable field. But can this finding be generalized? Does the survival of any new word, in general, depend on the stability of its contextual field?

We examine the causal effect of field instability on word survival by employing dynamic Poisson mixed-effects models. Estimates with three different covariate specifications are reported in Table 2.1. Model (1) is a baseline model with only temporal dependence and fixed and random effects. Our focal variable, field instability, is included in Model (2). k-nearest neighbors, with $k = 50$, are used in the measurement of field instability. Model estimates with

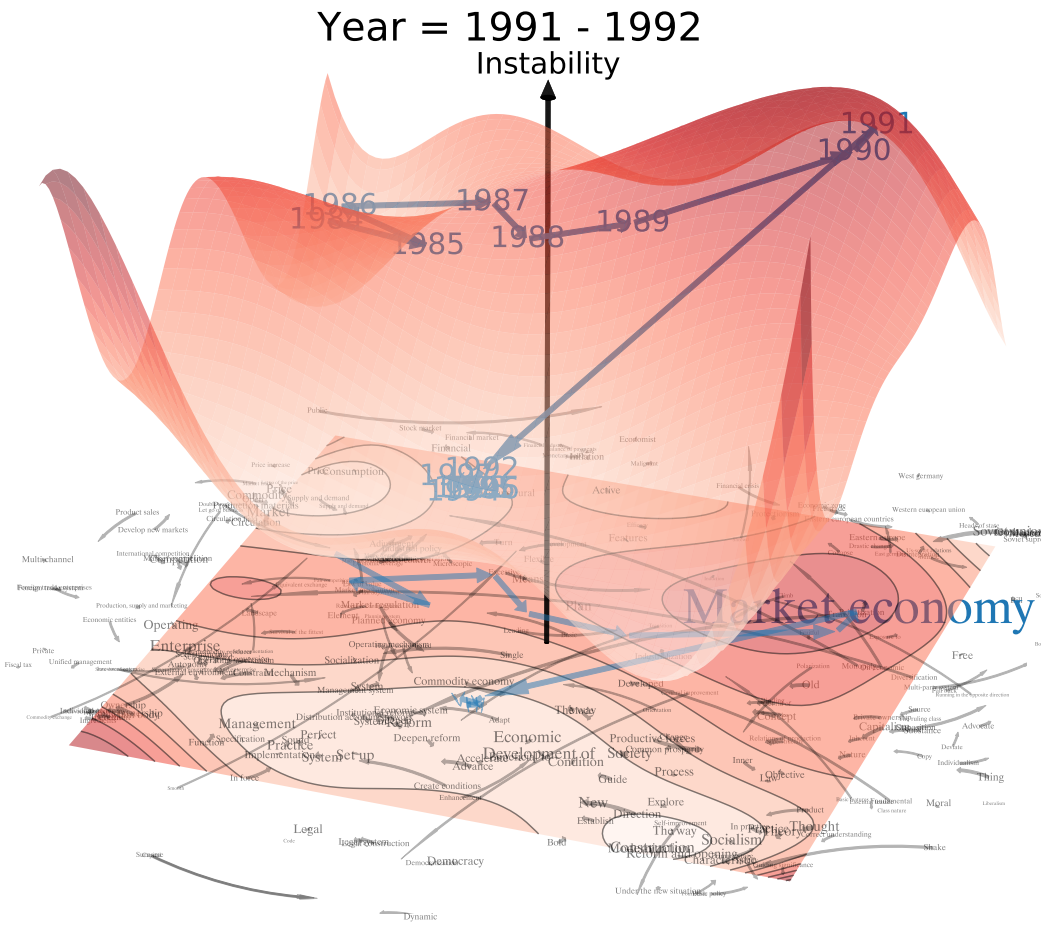

Figure 2.6: The xy-plane is a t-SNE (van der Maaten and Hinton 2008) visualization of the vector space around the word “market economy" from year 1991 to 1992. The z-axis represents the instability of the field. A polynomial surface of degree 6 was estimated based on the instability measures of the words involved. It was plotted along the z-axis and also projected onto the xy-plane. Changes in word's positions from 1991 to 1992 are indicated by the grey arrows. The blue arrows (on both the polynomial surface and xy-plane) trace the movement of the main phrase "market economy" since its inception in 1984. See Supplementary Section 7 for separate views.

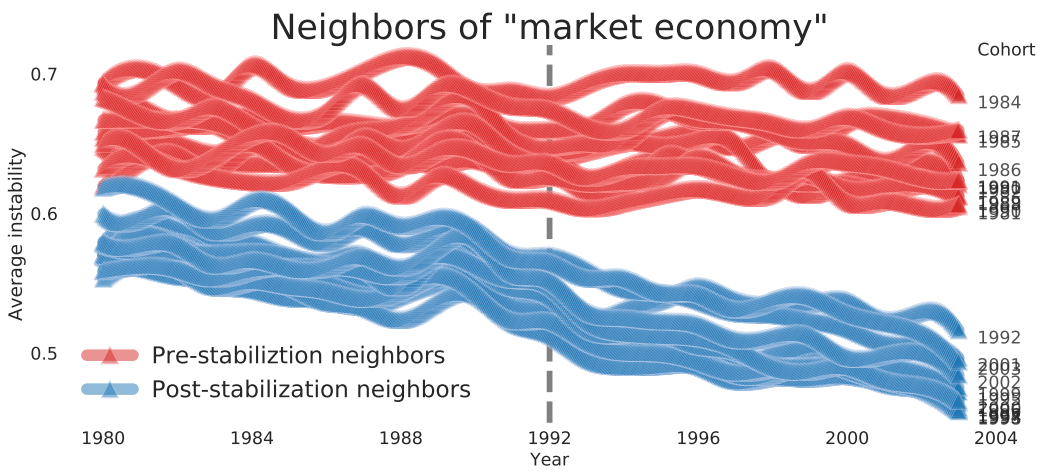

Figure 2.7: The yearly cohorts of the 100-nearest neighbors of "market economy" were first obtained from the word-embedding models, and their yearly average instabilities were plotted over time. It can be seen that the post-1992 neighbors were already more stable than pre-1992 neighbors before the year of stabilization. In other words, the phrase was settled in an existing stable neighborhood.

$k = 30$ and/or a different measure of field instability are reported in Supplementary Section 5. Self-instability is also controlled because it could be correlated with field instability, and previous studies have found that frequency and self-instability are highly correlated (Hamilton, Leskovec, and Jurafsky 2016b, 2016a). The table shows that Model (2) fits the data much better than the baseline model, and that field instability exerts a strong and direct negative effect on word frequency. In other words, the more unstable a word's field is, the less likely it is to survive. A permutation-based significance test is reported in Supplementary Section 5. Because the causal direction between self-instability and self-frequency is unclear, to avoid collider bias, only lagged self-instability is controlled in Model (3), and its effect is found to be minimal; thus, we could exclude self-instability as a confounder.

To investigate further whether the field effect is truly exogenous[20], we focus on the initial conditions of all the words during the first 12 months since their inception, and regress every

20. Here, we use the word "exogenous" to mean that the treatment variable is not correlated with the error terms. However, within our theoretical framework, the field effect offers an internalist (endogenous) explanation of rhetorical change.

Table 2.1: Mixed-effect Poisson regression (1-3) and logistic regression

Outcome	Frequency			Birth
	(1)	(2)	(3)	
Lagged log frequency	0.382^{***}	0.085^{***}	0.409^{***}	
	(0.004)	(0.005)	(0.007)	
Self-instability		-3.340^{***}		
		(0.006)		
Lagged self-instability			0.268^{***}	
			(0.006)	
Field instability		-5.600^{***}	-6.100^{***}	6.522^{***}
		(0.011)	(0.012)	(0.092)
Initial year ×	-0.067^{***}	0.062^{***}		
Lagged log frequency	(0.010)	(0.009)		
Constant	2.065^{***}	9.200^{***}	6.154^{***}	11.149
	(0.035)	(0.040)	(0.030)	(25.780)
Age fixed effects	Yes	Yes	Yes	
Year fixed effects				Yes
Variance of random effects				
Word	0.279	0.312	0.191	
Year	0.035	0.045	0.026	
Observations	149,775	149,775	110,613	260,153
Log-likelihood	$-2,123,668$	$-1,680,292$	$-1,278,355$	$-123,478$
Log-likelihood ratio test		$886,753^{***}$		
Bayesian Inf. Crit.	4,247,802	3,361,073	2,557,163	247,417

Note: Standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

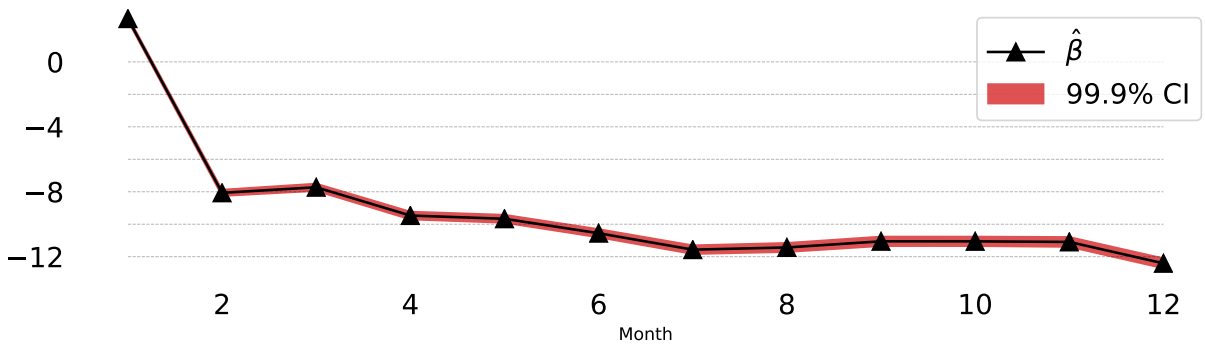

Figure 2.8: Effect of field instability on frequency in the first 12 months since inception

month's frequency on field instability, controlling for birth cohort in a Poisson regression. Coefficient estimates are plotted in Figure 2.8. The results show that the effect is positive in the 1st month and only becomes negative after the 2nd month. In other words, the words that are born in unstable fields have higher initial frequencies than those born in stable fields. The field stabilizing effect only comes later after birth. Year-by-year results are shown in Supplementary Section 6. The pattern is consistent across all years.

Similarly, when the question of where new words are born is asked, field instability is found to be positively related to birth. Logistic regression results are reported in the last column of Table 2.1 to show whether a word is more likely to be a new word or an old word given its field instability. In contrast to survival, birth is more likely to occur in unstable fields rather than stable fields.

In sum, the logic behind birth and survival is entirely different. New words are more likely to be born in unstable fields, and words that are born in unstable fields also have higher initial frequencies. However, only words that are later situated in stable fields are likely to survive.

Conclusion

As the CCP's mouthpiece, the *People's Daily* certainly reflects the Party leaders' will and is definitely influenced by outside forces on a daily basis. However, our analyses suggest that external shocks and the whims of power only add rhetorical ingredients. Long-term discourse cannot be easily changed at will. Self-consistency is always a fundamental principle that glues an ideological system together. As already noted by many China experts, the Chinese reform was not a pre-planned project, but always proceeded in a "snowballing" fashion (Shirk 1993). The CCP is known to be good at using old repertoires to promote new ideas (Perry 1980). Our analyses demonstrate again that path dependency is a central feature in cultural evolution.

What is striking about our results is that time-dependent patterns in the *People's Daily* were driven by a few time-independent rules. Although fundamental changes in the economic mode of production had happened in CCP China, discourse changes in the Party's mouthpiece had proceeded in a highly consistent fashion. External shocks as large as the 1989 Tiananmen crackdown could quickly die out in such a system and had little long-term effects on discourse change. In order to have long-term consequences, new elements must be internally incorporated into existing contexts. Controversial new concepts like "market economy" only became stabilized in such a system by being contextually attached to a safe harbor.

Our results also suggest that the birth and survival of novelty involve two different processes. Birth mostly comes out of the instability of a system, but a new element can also quickly die in an unstable field. In order to have a long-term impact, it is better to be integrated into a stable part of the system. The results are consistent with Padgett and Powell (2012)'s recent theory of invention. Similar rules have been previously found in knowledge production (Uzzi et al. 2013).

In a rhetorical system, words are not isolated "atoms", but are embedded in contexts.

The fate of every word – its survival, prospering, and death – depends on the context in which it is used. When a word is interlocked with other words in a stable context, a positive feedback loop can be activated to generate more re-occurrences of it (Padgett and Powell 2012). In the long run, changes that can reproduce survive. In this sense, an ideological system has an autonomous rule of its own.

Discussion

Could it be the case the internalist patterns that were found in this study were only epiphenomenal? In other words, could they merely be shadows of some material activities rather than having autonomy of their own? Our field explanation is a meso-level explanation. The micro-level actions that determined the movement of the words were not observed.

We, however, do not naively assume that words can move by themselves. Words, of course, are always moved by people, and their meanings oftentimes are arbitrarily decided by people. A full explanation of their movements would require an understanding of the actions behind their usages. However, words also only become meaningful when they are organized in some symbolic structures, and the symbolic structures could exist independent of their authors' material interests.

Because our field analysis provides a parsimonious explanation of the movements of words, in light of Occam's razor, we argue that the patterns we observed were not epiphenomenal. If the patterns were merely reflective of some external phenomena, it would be very challenging to offer an externalist explanation that is as parsimonious in explaining the complex economic and political reality that the texts represent. For instance, in the case of "market economy," the phrase was first used in the newspaper for seven years without a stable contextual usage. It acted like a free-floating signifier without any clear external referent. After it settled down in 1992, it became an anchor in the field. A closer look at its final neighborhood would reveal that its neighboring words are anchors words like "reform," "socialism," and "economic

development" as well. Rather than having any clear external referents, those words acted more like "coded symbols" that only make other words in the surrounding space meaningful. Although there was certainly an external force in 1992 that triggered the settlement of "market economy," the explanation that our field analysis offered is still internalist in the sense that the external push only had an effect insofar as it was internalized in the rhetorical field. Without the parsimonious explanation we offered, it would be very challenging to link the bizarre movement of the phrase to external events in the material world.

Marx was famous for inverting the Hegelian dialectics. In this study, we attempted a second inversion and applied a Hegelian analysis to a Marxist discourse. However, our attempt is very modest. We do not seek to provide a comprehensive causal understanding of the role of ideology in the development of history. Our goal is only to demonstrate, with causal precision, that there is some internal autonomy in the development of ideas. We do not deny that ideology is constantly influenced by material interests and external forces. But using Bourdieu's "game" metaphor, we would like to argue that even if ideology is just a game, it is nevertheless a game to be played, and in its playing, the game acquires a life of its own.

Our findings are also related to the classic question of "where novelty comes from." In studies of knowledge production, it is well-known wisdom that successful inventions are made out of atypical combinations of old elements. Actors, organizations, and ideas that occupy "structural-hole" positions enjoy a relative competitive advantage over their peers (Burt 1992). Many existing studies were performed on cross-sectional data. In cross-sectional data, failed cases do not get observed. Many cross-sectional analyses might suffer from survival bias. Our findings are partly consistent with conventional wisdom insofar as we found that novel words emerged in unstable contexts. However, our findings also suggest that the problem of survival might have been overlooked in previous studies. Future studies are needed to clarify the distinction between the two processes.

Our findings are generally consistent with either a Weberian explanation or a Durkheimian explanation. Field stabilization could arise out of rationalization or some symbolic structures. Our empirical analysis does not provide a way to test which one is true. Future work will be needed to answer our question with more precision.

Traditionally, the meaning of texts can only be qualitatively interpreted by social science experts in *verstehen* modes, the method of which is limited by both the volume of texts that an expert can read and the reader's subjective biases. With the help of machine learning, the methods employed in our research provide new quantitative ways for extracting the most salient internal patterns from massive textual data, which can be applied to many other areas of research in the age of information overload.

Methodological Appendix

The probability distribution of words between two time periods was compared by calculating the Kullback-Leibler (KL) divergence, defined as:

$$D_{\mathrm{KL}}(P\|Q) = \sum_i P(w_i) \log \frac{P(w_i)}{Q(w_i)} \tag{2.1}$$

where P and Q are probability distributions of words in two different time periods, and $P(w_i)$ and $Q(w_i)$ are the probabilities of word i under P and Q, respectively. The KL divergence is a precise measure of the extra bits of information that are needed when Q is used to approximate P (Kullback and Leibler 1951). A time series $\mathbf{d}$ can then be constructed with:

$$d_t = D_{\mathrm{KL}}(P_t\|Q_t), \tag{2.2}$$

where P_t is the probability distribution of words in the tth time period and Q_t is a prior reference distribution. The method allowed us to quantify the speed of change through time.

We embedded the words from every year into a 400-dimensional vector space. Specifically, we used the skip-gram model. Given a sequence of training words $w_1, \ldots, w_N$, the objective of the skip-gram model is to find a two-layer representation of words to maximize the average log probability,

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{j=-k}^{k}\log p(w_{i+j}|w_i) \tag{2.3}$$

where k is a window size. The probability of correctly predicting the word w_i given the word w_j is defined as

$$p(w_i|w_j) = \frac{\exp(\mathbf{u}_{w_i}^{\top}\mathbf{v}_{w_j})}{\sum_{l=1}^{\mathcal{V}}\exp(\mathbf{u}_{w_l}^{\top}\mathbf{v}_{w_j})}, \tag{2.4}$$

where $v_{w_j}, u_{w_i} \in \mathbb{R}^d$ are the input and output vector representations of the word w_j and w_i, and $\mathcal{V}$ is the set of all unique vocabularies in the corpus (Mikolov, Sutskever, et al. 2013). Huffman Trees were used to speed up computation, and optimal solutions were found through stochastic gradient descent. Equations [3] and [4] made it possible to extrapolate the probability of the co-occurrence of a set of words given a year's learned vector space by utilizing both the input and output representations of the words (Taddy 2015).

Given two word embeddings, the Procrustes Analysis finds the best orthogonal linear transformation (reflection and/or rotation) of one with respect to the other (See Supplementary Section 4). It allowed us finding the (in)stability of a word w from year $t-1$ to t as $\cos\theta_{w,t}$. An alternative measure is the substitution rate of a word's k-nearest neighbors (Palla, Barabási, and Vicsek 2007) (See Supplementary Section 5). We used both methods. These methods allowed us to measure the instability of the contextual field the word was in, as $\frac{1}{k}\sum_{j=1}^{k}\cos\theta_{w_j,t}$ where w_j's are w's k-nearest neighbors.

To study the causal effect of field instability on word survival, we considered the following

dynamic Poisson mixed-effects model:

$$\begin{aligned} \log(\lambda_{i,t}) &= \alpha + \gamma \log F_{i,t-1} + \mathbf{x}_{i,t}^{\top}\boldsymbol{\beta} + \mu_i + \nu_t \\ F_{i,t} &\sim \text{Poisson}(\lambda_{i,t}), \end{aligned} \tag{2.5}$$

where $F_{i,t}$ is the frequency of word i in year t, $\mathbf{x_{i,t}}$ are covariates of interest, and μ_i and ν_t are the random effects of word i and year t. A lagged log frequency term was included to account for temporal dependence, and the yearly frequencies of every new word since the first year after its inception were included in our analysis. (The definition of new word is explained in Supplementary Section 5.) To model initial conditions, we assumed that the first month's frequency of every word was truly exogenous and used that as $F_{i,t-1}$ when $t-1$ is the year of inception. For that reason, an "initial year $\times$ lagged log frequency" interaction term was also included to account for a different γ. The fixed effect of age (t - birth year) was included as a covariate. This Poisson model is equivalent to an event history model, assuming piece-wise constant hazard rates in which the outcome variable is the re-occurrences of words.

Availability of Data and Code

All necessary data and codes for replicating the findings of this study are publicly available at https://drive.google.com/file/d/1gkihvxJ24mhB6qasU635HdprfqezFSXO/view?usp=sharing.

CHAPTER 3

MAPPING INTERNATIONAL CULTURE: PERCEPTION OF IDENTITIES IN A HUNDRED YEARS OF GOOGLE BOOKS

Shilin Jia

Chapter Abstract In this chapter, we propose a methodological framework for measuring how countries historically perceived other identities when they engaged in war and trade with others. The method relies on the use of n-grams in millions of digitized books that Google has scanned since 2009. Based on word co-occurrences in the Google Ngrams, we are able to train yearly neural-probabilistic word embeddings from the corpora. The word-embeddings allow us to extract meaningful dimensions from the space and map different identities to the dimensions. This enables us to take a dive into history and understand how different language communities perceived other identities. We demonstrate reasonable success in validating our measures with external measures on international war and trade. However, despite our success in obtaining promising results in American English, British English, and French, further effort is still needed for constructing good measurements in German, Italian, and Russian.

Introduction

With the development of social surveys, social scientists have been increasingly able to gauge the minds of the public and understand how the public thinks about and perceives various social and political issues. One topic of interest that frequently makes newspaper headlines is how the general public in a nation perceives foreign countries in the world. Feeling thermometers have been widely used in social surveys for answering such questions. The feeling thermometer asks respondents how favorable their views are on a certain country,

ranging from "cold" to "hot." In the United States, Gallup and the Chicago Council on Global Affairs have been including feeling thermometers in their national public opinion surveys since the 1970s (Schneider 1985). Since the early 2000s, Pew Research Center has been conducting a multi-national survey asking respondents in dozens of countries about their perceptions of foreign countries (X. Li 2021). Such longitudinal records allow researchers to develop a historical understanding of how the general public in a country changes their perceptions in response to external events. For instance, Pew's surveys show that Americans' perception of China has reached a historical low following the U.S. and China's recent clash in trade wars and on the spread of coronavirus (X. Li 2021; Pew Research Center 2020). Scholars found that public perceptions are correlated with foreign policies (Lee and Hong 2012). In democratic countries, such as the United States, change in public perception oftentimes even precedes change in national foreign policies (Page and Shapiro 1983). Studies on public perceptions could yield valuable insight into understanding international relations and provide explanations such as why democratic couturiers don't fight each other (Gries et al. 2020). However, because those surveys have been only been systematically conducted in recent times, how people in different nations historically perceived other nations remains a relatively unknown topic. It is impossible to travel back in time and ask people in the past about their perceptions. Also, perceptions of other identities could be inherently multi-dimensional, but a single-measure feeling thermometer ranging from "cold" to "hot" could be misleading (X. Li 2021). Scholars' understandings of this topic have so far still been relatively limited by the availability of data.

One way to measure human biases is through language (Dodds et al. 2015; Kozlowski, Taddy, and Evans 2019). Instead of asking people about their views on certain subjects, researchers can observe what people say or write in their languages, and more specifically, in what contexts subjects of interest are mentioned. In this chapter, we propose a new research framework for understanding how literate elites in different language communities

historically perceived other foreign identities. We use the Google Ngram Corpus to build yearly word-embedding spaces in 6 languages that are currently available. Then by projecting foreign national identities onto two in-group vs. out-group dimensions, we are able to obtain historical measures of foreign identities in these two dimensions. We show that in American English, British English, and French, there is consistent change in perceptions of other identities when their countries are at war with other countries.

Measuring Cultural Dimensions in Word Embeddings

Recent development in neural probabilistic language models has provided social scientists a way to map millions of words and documents from a gigantic corpus into a high-dimensional vector space (Bengio et al. 2003). Theoretically, this modeling approach resonates with structuralist theories and the distributional hypothesis in linguistics (Saussure 2011; Harris 1954; Firth 1957). The theory holds that the meaning of a word in a language is defined by its surrounding words, and words that appear in similar contexts have similar meanings. Neural probabilistic language models represent each word in a corpus as a vector and use the vectors to optimize the task of predicting words given their contexts. The vector representations, as by-products of the models, turn out to have high interpretative values. The models are able to map words that occur in similar contexts into proximate locations in a vector space.

The Word2Vec model and its skip-gram variant developed by Mikolov et al. has become a state-of-the-art technique for building vector representations of words from a large corpus (Mikolov, Sutskever, et al. 2013; Mikolov, Chen, et al. 2013). Based on a simplified two-layer neural network design, it is able to process huge volumes of texts fastly and efficiently. Computational linguists have utilized the model to study change of words' meaning over time (Hamilton, Leskovec, and Jurafsky 2016a).

Interestingly, computational linguists have also found that word-embedding models are not only able to map local clustering of words with similar meanings but are also capable of

performing more complicated language tasks such as solving analogy tests. The models are able to solve standardized test questions in the form of "a is to b as c is to __" (Rumelhart and Abrahamson 1973; Mikolov, Sutskever, et al. 2013). A classic example is the question "man is to woman as king is to __." Word-embedding models trained on a reasonably large English corpus are usually able to yield $\overrightarrow{king} + \overrightarrow{woman} - \overrightarrow{man} \approx \overrightarrow{queen}$.The result suggests that $\overrightarrow{man} - \overrightarrow{woman}$ and $\overrightarrow{king} - \overrightarrow{queen}$ are parallel directions in the vector space, and there might exist a global gender dimension in the space. The reason why such a dimension exists is that "man" and "woman" are similar words in the sense that they are both nouns used to designate genders. They are oftentimes used interchangeably in texts except for the crucial difference that "man" is used more often in masculine contexts, and "woman" is used more often in feminine contexts. Subtracting the vector representations cancels out their similarity and keeps their difference. And a similar difference is also kept in taking the subtraction of "king - queen." The models are able to solve many different kinds of semantic and syntactic analogy tests such as "France is to Paris as Italy is to Rome" and "bad is to worse as big is to bigger." They are also able to learn explicit and implicit biases and stereotypes in languages. For instance, Bolukbasi et al. (2016) find that world embedding models could yield parallelograms such as "man" is to "computer programmer" as "woman" is to "homemaker" and "father" is to "doctor" as "mother" is to "nurse." Caliskan, Bryson, and Narayanan (2017) find that in their word-embedding models, African-American names are more associated with unpleasant words than European American names.

Kozlowski, Taddy, and Evans (2019) propose a method to project cultural items onto cultural dimensions. The idea is that antagonist pairs such as $\overrightarrow{man} - \overrightarrow{woman}$ and $\overrightarrow{rich} - \overrightarrow{poor}$ represent some global directions in the vector space. And by projecting words such as "scientist", "nurse", "hamburger" and "wine" to those directions, researchers are able to tell how feminine/masculine and rich/poor each of these items is in a given linguistic community. Presumably, this measure would yield a higher loading of "king" and a lower loading of

"queen" in a feminine/masculine dimension. The measure allows researchers to map as many items as there are in a corpus to the same linguistic dimension and yield a single measure of how these items score in that dimension. In our study, we apply this method to study international cultural perceptions.

Research Design

Building yearly word embedding models from the Google Ngram corpus

We trained our word-embedding models from 6 languages of the Google Ngram corpus. The Google Ngram corpus is one of the largest publicly available corpora that contain a huge amount of texts produced in human history. It is produced out of Google's massive effort in digitizing historical texts into Google Books. Although the original texts are copyrighted, Google releases all of the ngrams that appear in the texts. An ngram is n-words co-occurring together in a text. The maximum ngram length that Google supports is 5. The corpus gives a count of the number of times each ngram appears in the Google Books published in a given year. All n-grams that appear at least 40 times across all years are included. Researchers have used the ngram counts to tell the relative importance of different cultural items, identities, and political figures in human history (Michel et al. 2011; Jiang, Xi, and Xie 2020). The following list is a snippet of 5-grams in UK English.

've a bloody good mind

've a decision to make

've a feeling that he

've a good mind ...

've a job for you

've a kind and loyal

've a legal right to

've a life to live

've a mind to show

've a mortal aversion to

've a notion in my

've a notion that if

've a plenty for them

've a pretty shrewd idea

The ngrams have also been used for learning language models (Hamilton, Leskovec, and Jurafsky 2016a). In this study, we use the same corpus that has been previously used by many researchers to build yearly word-embedding models from 1900 to 2019. We also extend our research scope to 6 languages including the US and UK English, French, German, Italian, and Russian. Google uses the location of publishers to distinguish between the US and UK English.

When processing a normal text, Word2Vec uses a sliding window to predict and learn what other words would appear in a word's context window. Training a Word2Vec model on Google 5-grams is as if we are processing a normal text through a sliding window of 5. Google does not provide information about what specific books each n-gram appears in, and each year's corpus is released as a whole. Scholars using Google Ngrams also noted that the composition of Google Books could have changed significantly during the last century (Kozlowski, Taddy, and Evans 2019). Around 2009, the primary source of Google Book also changed from books scanned from libraries to digitized materials released directly by publishers. Although the Google Books corpus is large, it certainly does not represent all speakers in a language community. It is also probably not a representative sample of all books published in a language. However, given its gigantic size, word-embedding models are able to learn common linguistic dimensions shared by authors in a language. The dimensions

can also tell common biases and stereotypes in a language community.

We sampled from all 5-grams to construct our yearly corpus and trained our yearly word-embedding models. Figure 3.1 shows the sizes of the 5-gram corpora across time. Because the corpus sizes are dramatically different across years, to make our yearly models comparable, we randomly sampled the same number of 5-grams from each year's corpus. The yearly sample size we used is the smallest yearly corpus size in the full corpus. For instance, in British English, there were least publications in 1944 during the Second World War period, and the yearly sample size we used for other years is equivalent to the size of the 1944 corpus. To account for randomness in the texts, we bootstrapped 20 samples for each year using sampling with replacement. The bootstrapped samples would allow us to construct confidence intervals for our time series. Sampling with replacement has the effect of dropping rare 5-grams from our yearly corpus. It helps us to focus on central tendencies in the corpus. The Russian language had a major language reform in 1918 following the Bolshevik Revolution. Many modern words did not exist in the Russian corpus prior to 1918. So we chose 1918 as the beginning year of our Russian models. Unfortunately, for US-English, the corpus size is so large that training yearly models according to our framework would cost a huge amount of time. We were not able to train all of our models in time. With a compromise, we trained one model for each year using 20% of the least corpus size. The 20% of the least corpus is still larger than the yearly sample sizes of all other languages. Presumably, measures in US English would have shorter confidence bands than the ones observed in other languages. Table 3.1 summarizes information about our samples.

After obtaining our yearly samples, we trained yearly word-embedding models using the skip-gram model (Mikolov, Chen, et al. 2013). With the implementation of a negative sampling framework, what the skip-gram model does, in a nutshell, is to use a logistic regression to distinguish word pairs that actually occur in the corpus and randomly generated "negative" word pairs that don't appear in the corpus. Through stochastic gradient descent,

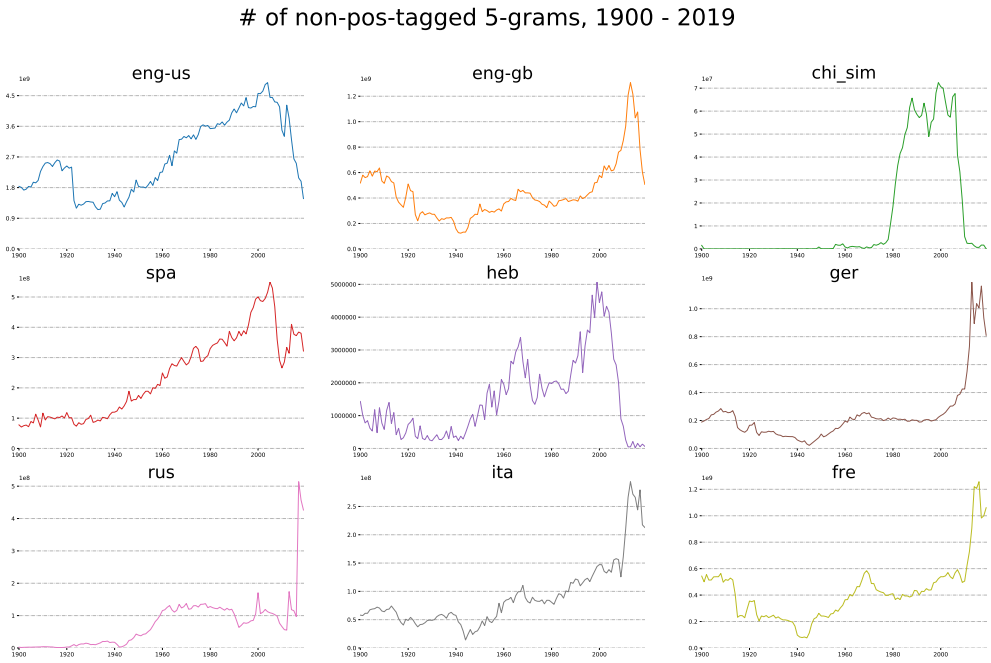

Figure 3.1: Total number of Google 5-grams in all available languages

Table 3.1: Sample sizes of our study

language	yearly sample size (total # of 5-grams)	bootstrapped samples	years
US English	2.3e8	1	1900-2019
UK English	1.2e8	20	1900- 2019
French	7.6e7	20	1900- 2019
German	2.1e7	20	1900- 2019
Italian	1.4e7	20	1900- 2019
Russian	1.4e6	20	1918- 2019

it finds the best vector representations of words to maximize the log-likelihood of the logistic regressions. More details about the inner working of the model are explained in the appendix of this chapter.

External data on international war and trade

We used data on international war and trade to externally validate the soundness of our cultural measures.[21] The international war data comes from the Correlates of War Project [22]. The project has records of all international wars between nations from 1816 to 2007. It records whether a country is an ally or enemy of another country in a specific year. The international trade data comes from the Historical Bilateral Trade and Gravity Data set (TRADHIST) constructed by Fouquin, Hugot, et al. (2016). The dataset contains import and export trade volumes between nations from 1827 to 2014. It also has each country's historical GDP. We are interested in exploring whether our cultural measures correlate with international trade and war.

In-group vs. out-group dimensions

To construct our cultural measures, we focus on two analytical dimensions in in-group vs. out-group distinction. The first dimension we focus on is the "friend-enemy" dimension. As theorized by German political philosopher Carl Schmitt (1976), the friend-enemy distinction is a fundamental distinction in politics, which specifies with whom a political entity should form an alliance and make enemies in political struggle. We expect the dimension to be salient in a language community during wartime.

To measure more implicit cultural distinctions, we constructed a separate "we-they" dimension. Based on the social identity theory proposed by Tajfel et al. (1979), a basic process

21. Special thanks to Tamara van der Does for compiling and formatting these data.

22. https://correlatesofwar.org/

in people's everyday socialization is distinguishing who are part of "them" and who are part of "us." In this process, people use stereotypes to create group images for "others" and "us." This process is an origin of inter-group conflict in social life. It is more cultural and implicit. Word Embedding models provide us a tool to measure which identities are part of "them" and part of "us" in a language community. Because countries are not in direct conflicts most of the time, we expect that this dimension is more salient in telling subtle country-to-country relations during peacetime. We also expect that the we-they dimension is associated with how often members in a community interact with members in other communities. Therefore, it might have a stronger relationship with international trade.

Following Kozlowski, Taddy, and Evans (2019)'s method, for each yearly model, the dimension $\boldsymbol{d}$ is constructed as

$$\boldsymbol{d} = \sum_{w_i \in \text{positive-words}} \text{norm}(\boldsymbol{v}_{w_i}) - \sum_{w_j \in \text{negative-words}} \text{norm}(\boldsymbol{v}_{w_j}) \tag{3.6}$$

, and the loading of each selected identity-word in the dimension is the cosine similarity between its vector representation and $\boldsymbol{d}$.

Table gives the lists of positive and negative words we used in constructing the friend-enemy and we-they dimensions in 6 languages. Figure 3.2 shows that the chosen pairs in UK English are indeed parallel throughout history, at least in 2-dimensional visualizations. The visualization suggests that the dimensions exist.

Results

Time series plots

After constructing the yearly in-group vs. out-group dimensions, we were able to project selected national identities to each dimension and make time series plots.

Table 3.2: Postive and negative words used for constructing dimensions

language	dimension	positive words	negative words
US/UK English	friend-enemy	friend, ally	enemy, foe
US/UK English	we-they	we, us, our, ours, ourselves	they, them, their, theirs, themselves
French	friend-enemy	ami, amis, amie, amies	ennemi, ennemis, ennemie, ennemis
French	we-they	nous, on, notre, nos	ils, eux, leur, leurs
German	friend-enemy	freundin, freund	feind
German	we-they	wir	sie
Italian	friend-enemy	amica, amico	nemica, nemico
Italian	we-they	noi	esse, essi
Russian	friend-enemy	друзья, друзей, друзьям, друзей	враги, врагов, врагам, врагов
Russian	we-they	Мы, Нас, Нам, Нами, Нас	Они, Их, Им, Ими, Них

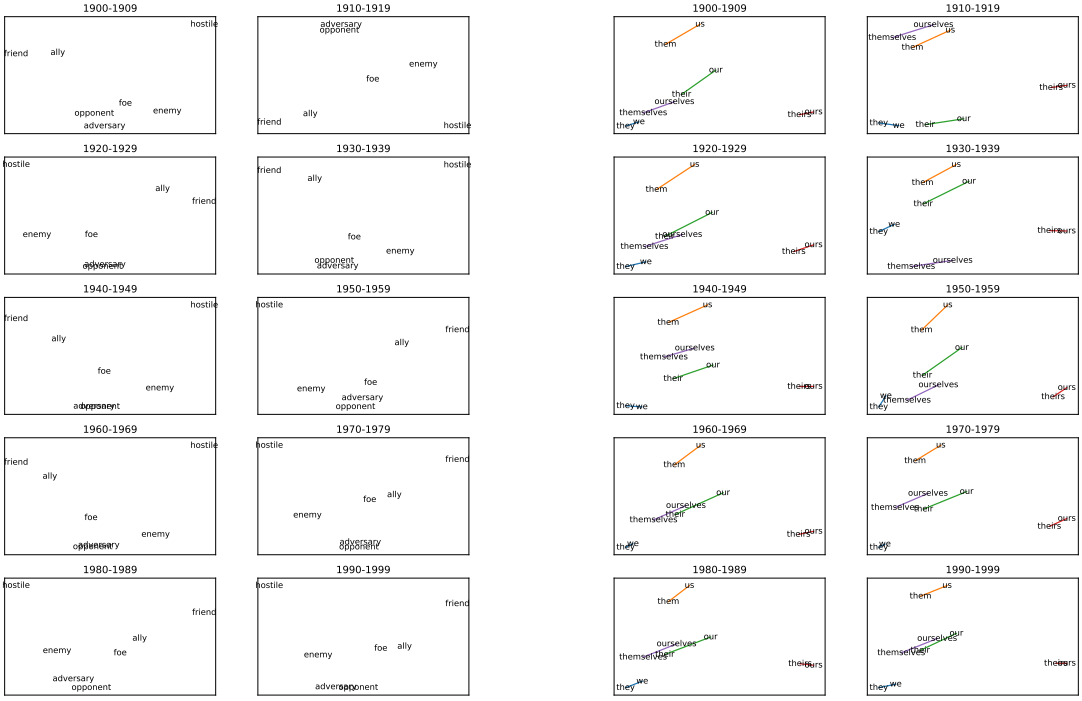

Figure 3.2: Friend-enemy and we-they pairs projected to 2D PCA dimensions

US-English

Figure 3.3 contains plots of selected national identities in the friend-enemy dimension in US-English. Without external validation, the plots make intuitive sense. Unsurprisingly, USA and Canada are the most friendly nations in American English. The two world wars had significant impacts on the friendliness of Germany, and Germany remained the least friendly western European country most of the time. Italy also experienced a similar drop during WWII but not WWI. Among Asian nations, Japan experienced a significant drop during WWII while China experienced a considerable boost. China and Japan's (as well as India's) friendliness scores were almost in parallel up till the outbreak of WWII. They also began to become indistinguishable from each other since the mid-1960s. Korea experienced a significant drop during the Korean War. Philippines as a former U.S. colony experienced a significant drop during WWII probably because of its independence movement. Vietnam experienced a significant drop around the time of the Vietnam War, but it bounced back after the 1970s and became indistinguishable from the Philippines. Thailand, on the other hand, has historically been a neutral country and is the most friendly Asian country most of the time. It nevertheless experienced a huge dip in WWII probably because it joined Japan's alliance during the middle of the war. Among Latin American countries, Mexico has always been the most friendly nation. Cuba experienced a huge drop around the time of the Cuban Revolution and remained low after that. Among countries in the Middle East, many countries including Iran, Afghanistan, and Libya experienced a significant drop around the time of the Islamic Revolution. Iraq continued its drop till the First Gulf War. It slightly bounded back after the war but dropped again around the time of the Second Gulf War. Lastly among Eastern European countries, the word "Soviet" has always been a highly hostile term. Its score started to decline even before the end of WWII. The image of Russia, however, was separate from that of the "Soviet." After the end of the Cold War, all former communist countries (except for Yugoslavia, which ceased its existence) seem to have

enjoyed an increase. Overall, most of the time series conform to conventional understandings of the United States' foreign relations. Wars seem to have played a major role in influencing other nations' perceived friendliness in US English. It is also worth noticing that countries in the same regions seem to oftentimes co-vary a lot. However, there does not seem to be any universal co-variation among all countries.

Figure 3.4 are the corresponding plots in the we-they dimension. War does not seem to play as much influence in those time series. And the graphs seem to be less interpretable in terms of major international events. For instance, although Germany was an enemy during the two world wars, its we-ness is not distinguishable from many other Western European nations including even America itself during wartime. Interestingly, although "UK" was a relatively novel abbreviation used in the history of American English, it became one of the most "us" nations in the latter half of the 21st century. However, "Britain," as an old word, remained the lowest "us" nation. The difference could tell some path dependency in the use of language. Among Asian nations, Japan also didn't experience a significant drop during WWII but had been in steady decrease till the end of the 1970s. After that, its we-ness started to bounce back along with China and Korea. Thailand, although is one of the most friendly Asian countries in US English, isn't as "us" as China, Japan, and Korea. More gradual and long-term international interactions in trade and immigration perhaps could explain the difference. Among Middle Eastern countries, Iraq also didn't experience a dramatic drop during the Gulf Wars. Lastly among Eastern European identities, "Russia" and "Soviet" started their great diverge since the middle of the 1980s probably because many Russians immigrated to the United States, and Russia itself also became a non-communist country. Overall, only a few well-known geopolitical crises, such as the Vietnam War, the Cuban missile crisis, and the Iranian Revolution seemed to have played a role in influencing the "us-ness" of other national identities. The we-they dimension is more cultural than political and is less reflective of high-level politics. Interestingly, although Mexico is the

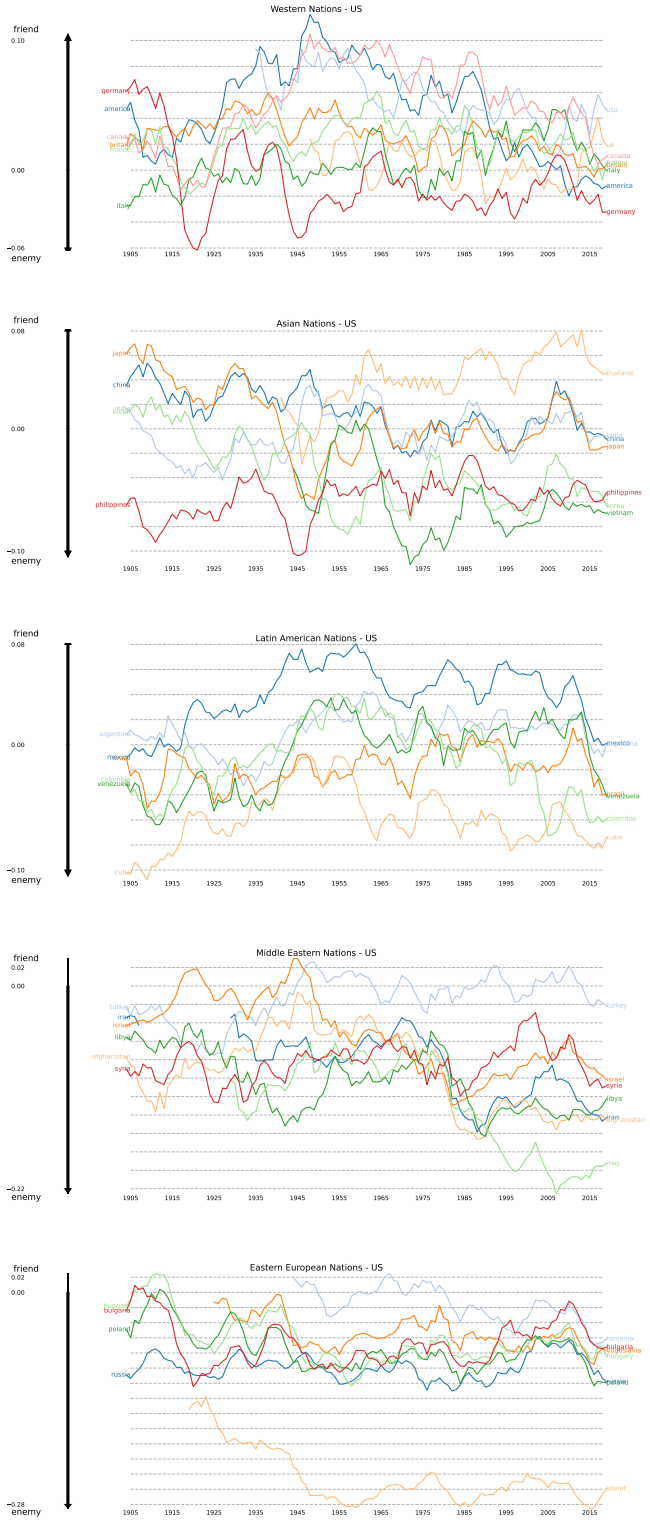

Figure 3.3: Loadings of selected countries in the friend-enemy dimension in US English

most friendly Latin American nation in US English, it is the least "us" identity in the we-they dimension. This suggests that frequent interactions do not necessarily lead to more incorporation. On the political level, Mexico has always been a close ally of the United States. But on the cultural level, its identity has always been more treated as "others" than "us."

UK-English

Figure 3.5 and 3.6 are our results in UK English. Because we had 20 bootstrapped samples, we were able to create 90% confidence intervals for our measures. The confidence intervals help to tell that a lot of times, observed differences in the means are not statistically significant. There is a considerable degree of randomness in countries' loadings, which could depend on which 5-grams were drawn into our samples. However, the graphs are still able to pronounce some major historical changes. "America" and "USA" have always been the most friendly Western European identities in British English. They also seemed to enjoy some special favor during WWII and the beginning of the Cold War. Germany and Italy experienced significant drops during WWII. France also seemed to have experienced a drop during WWII probably due to its surrender. However, during WWI, as allies of the UK, France and Italy both experienced a boost contrary to Germany. Japan experienced a huge drop amid WWII. Thailand is again relatively always a more friendly nation. Burma, as a former British colony, enjoyed the same rank as India before WWII. However, probably because it severed its tie with the British Empire after gaining independence, its friendliness continued to drop in the second half of the 20th century. Among Latin American countries, Cuba clearly stands out as an outlier after the Cuban Revolution. Among Middle Eastern countries, Iran, Libya, Afghanistan, and Iraq followed similar paths around the middle of the 20th century with Iraq dropping to the lowest during the time of the Gulf Wars. Loadings in the we-they dimension also in general tell a different story than the friend-enemy dimension.

Figure 3.4: Loadings of selected countries in the we-they dimension in US English

Germany has always been more "us" than "they." Japan and China are the most "us" Asian countries. Latin American and East European countries are mostly indistinguishable from each other. "Russia" and "Soviet" experienced the same divide since the end of the 1980s. Again, countries in the same regions seem to co-vary. But there does not seem to be any global co-variation. We also projected the constituent nations of the United Kingdom to the we-they dimension. Interestingly, the use of "UK" seems to generally complement the use of "Britain," "England," "Scotland," and "Wales." When "we" become more "UK," "we" become less "Britain," "England," "Scotland," and "Wales." All the sub-national identities seemed to be in decline following the world wars. However, the trend has been reversed since the 1980s.

French

The French corpus is the third largest corpus that we have. Figure 3.7 and 3.8 are our results in French. The stories that the graphs tell are very similar to the US and UK stories with some French ingredients. The two world wars are clearly pronounced in the friend-enemy plots. Although France was perhaps seen as a German collaborator in the eyes of the British, it sees Germany majorly as an existential threat in its own language. Japan and Italy were both seen as enemies during WWII. Canada is unsurprisingly the most friendly Western European nation probably due to its close linguistic tie with France. Vietnam's friendliness dropped significantly after it claimed its independence. However, the image of "Indochine" remained relatively unchanged. Vietnam reverted back to the level of Indochina in the late 1980s, and these two identities became indistinguishable again as Vietnam gradually re-opened and rebuilt its tie with its former ruler. To French authors, although Vietnam was lost, the image of Indochina remains in their imaginary. The dropping of the image of Cuba and the deterioration of the relationships with the Muslim world seem to be universal across American English, British English, and French. Interestingly, the friendliness of Algeria started to deteriorate before the general decline of the friendliness of the Muslim world.

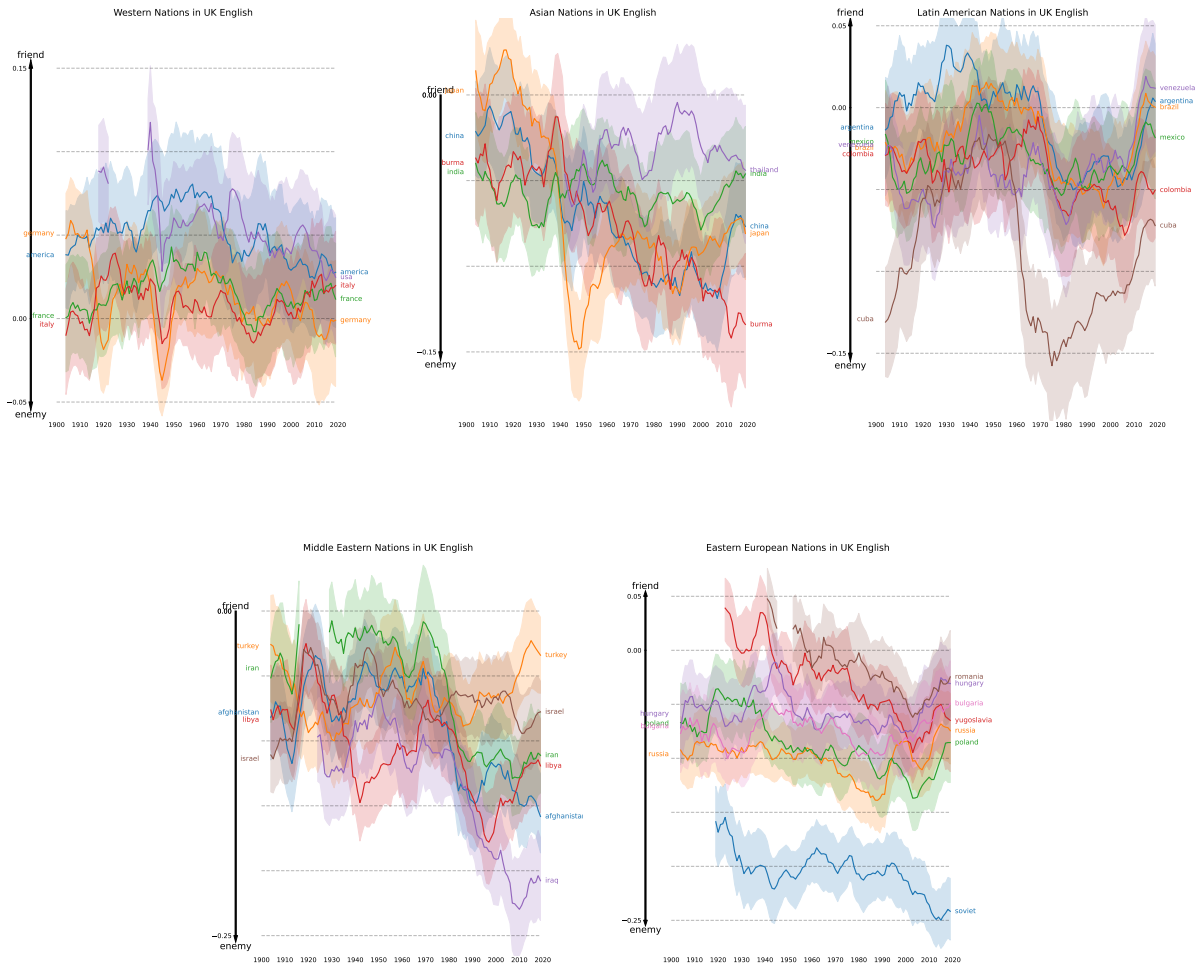

Figure 3.5: Loadings of selected countries in the friend-enemy dimension in UK English

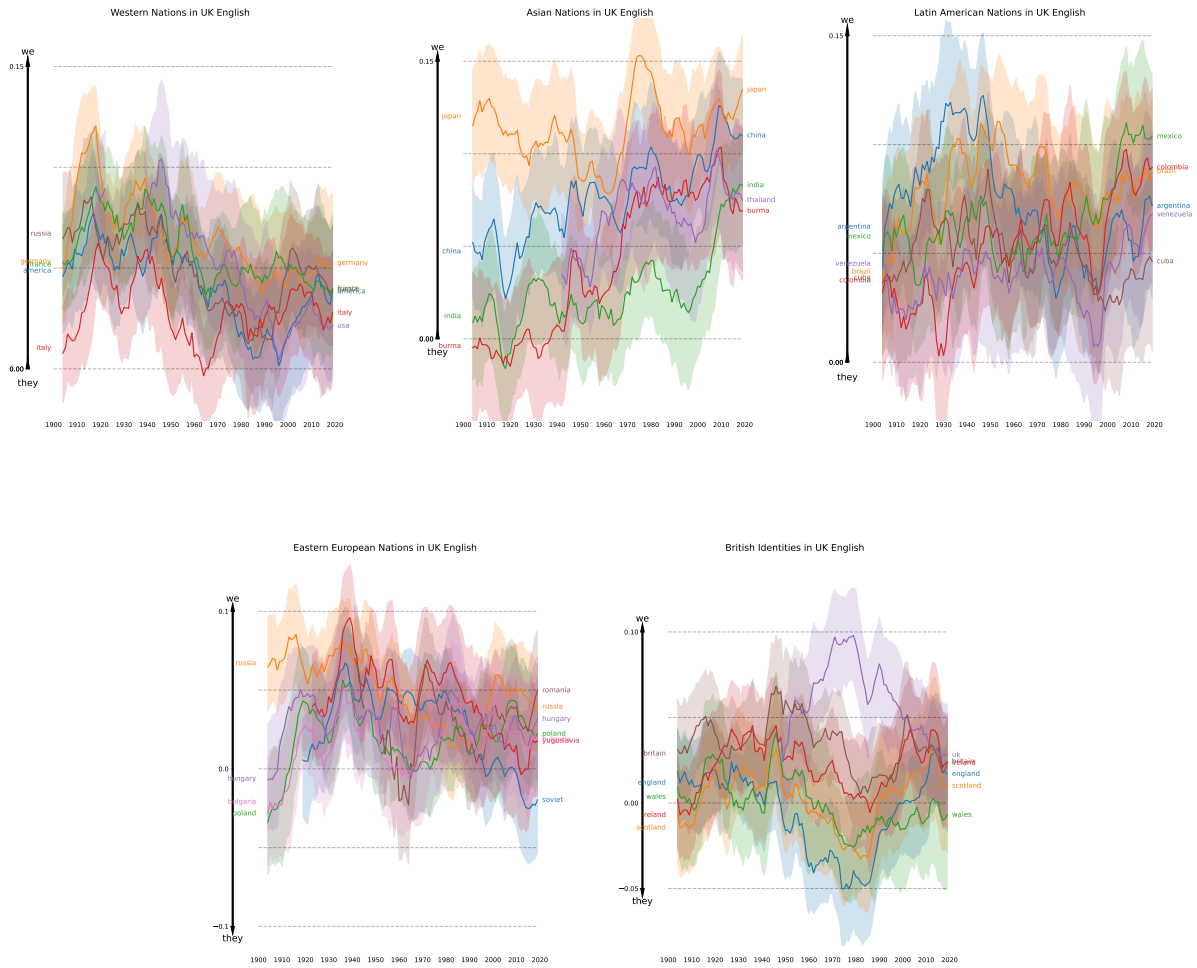

Figure 3.6: Loadings of selected countries in the we-they dimension in UK English

The decline should correspond to the Algerian Independence War. "Soviet" is also the least friendly identity that is only comparable to that of Germany during world wars with Russia enjoying a different path. In the cultural dimension, we also observe some similar general trends and some special French flavors. For instance, although Britain was politically an ally during world wars, it has always been more "them" than "us" in French publications. The friendliness of Indochina although remained largely unaffected by the independence of Vietnam, there is a clear sign that Indochina, which was one of the most "us-ish" identities at the beginning of the century, is no longer part of "us." The same happened to Algeria.

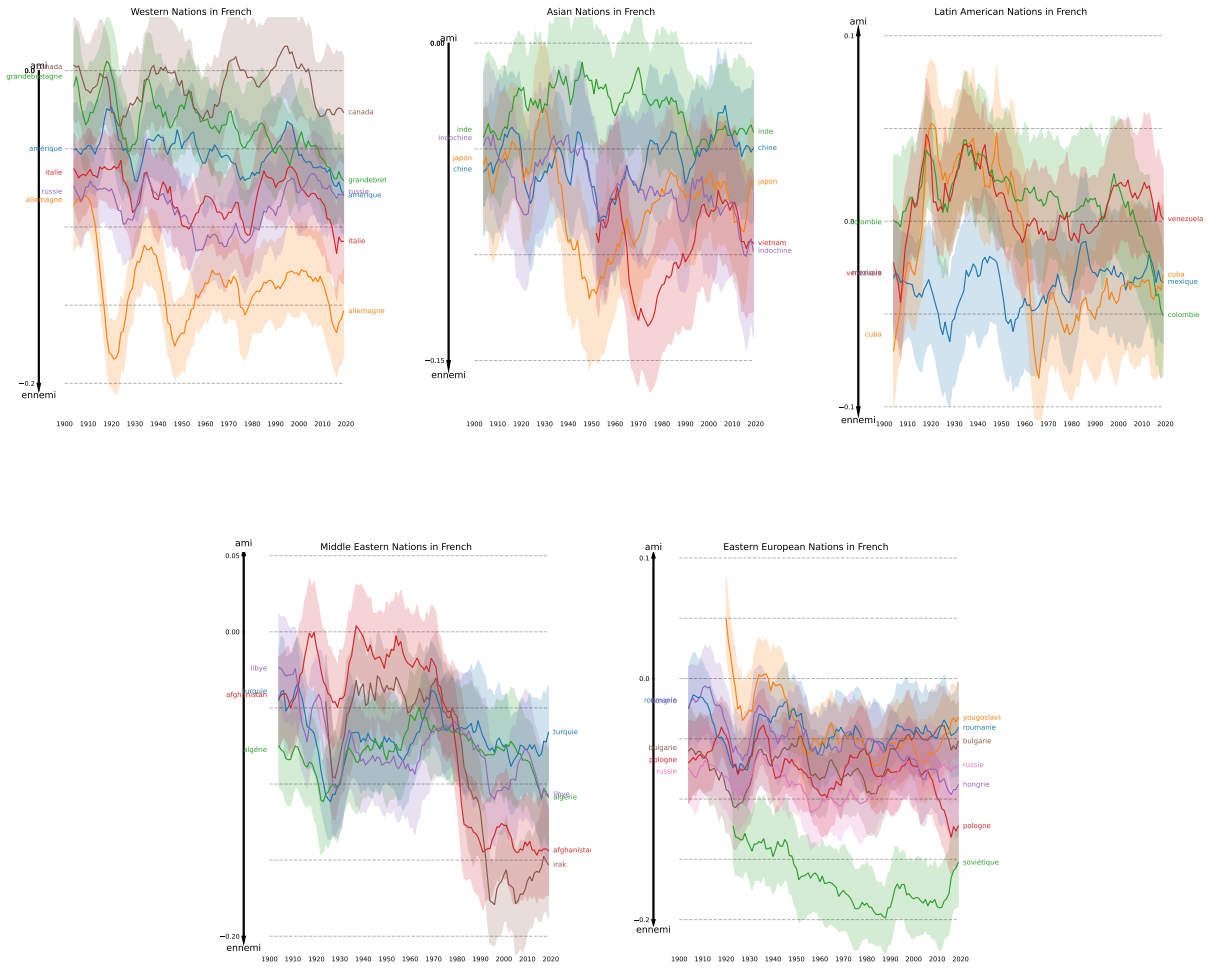

Figure 3.7: Loadings of selected countries in the friend-enemy dimension in French

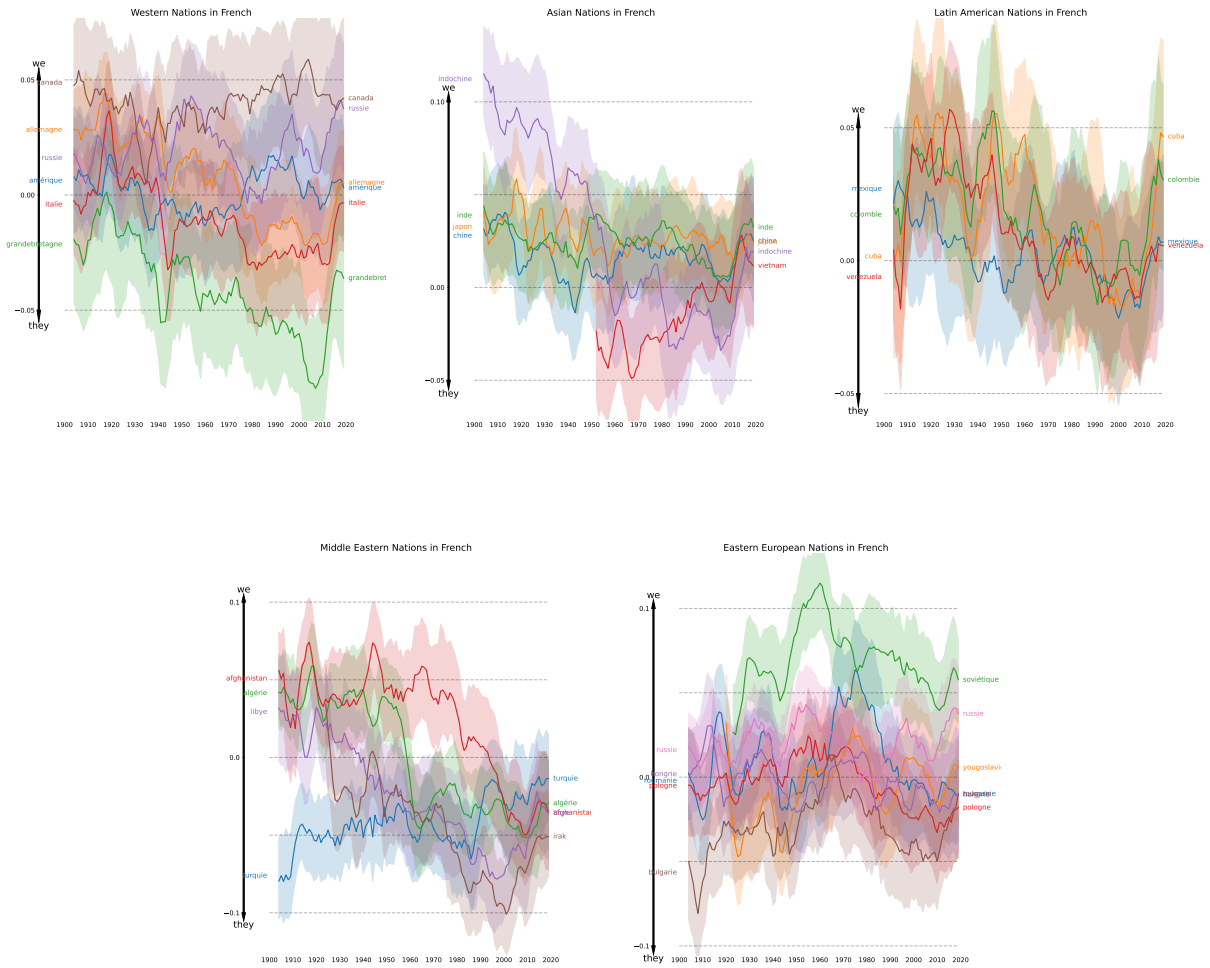

Figure 3.8: Loadings of selected countries in the we-they dimension in French

Less desirable results in German, Italian, and Russian

Unfortunately, we do not achieve the same level of interpretability in modeling German, Italian and Russian. Our results are shown in Figure 3.9-3.11. Some trends seem to be still interpretable. For instance, in German, there seems to be the same Clash of Civilization trend between the Western World and Muslim Countries. However, we will not go into detail in trying to interpret the graphs. Two basic problems observed in these graphs make us doubt the value of interpreting these graphs. First. there seems to be wide global co-variation among not just countries in the same region but all countries. This makes us wonder whether there is some spatial distortion in the word-embedding spaces. We observed similar spatial distortions when we first trained our English and French models. We were able to correct the problem by adopting a more methodologically consistent sampling scheme. (The details of our correction are documented in the Appendix.) However, even after applying the same sampling scheme to all corpora, we still observe spatial distortion in the word-embedding spaces of German, Italian and Russian. Because we are not able to exactly identify the cause of spatial distortion, we cannot tell whether the observed trends are due to some real historical processes or some noises introduced by our models. Second, some historical trends reflected in those graphs do not make intuitive sense. For instance, Japan and Italy were seen more as enemies rather than friends in German during WWII. The same is also true for Austria during WWI. We will discuss potential causes of the unfavorable results and try to continue to improve the quality of our models after this study.

Stretch of Dimensions during War Time

Lastly, one thing that we do observe across corpora is the stretch of dimensions during the world wars. Figure 3.12 plots the cosine similarities between positive words and negative words for all languages and all years. The higher the score is, the more the positive and negative words are used in dissimilar contexts. Except in US English, there is some significant

Figure 3.9: Loadings of selected countries in the friend-enemy and we-they dimensions in Italian

Figure 3.10: Loadings of selected countries in the friend-enemy and we-they dimensions in German

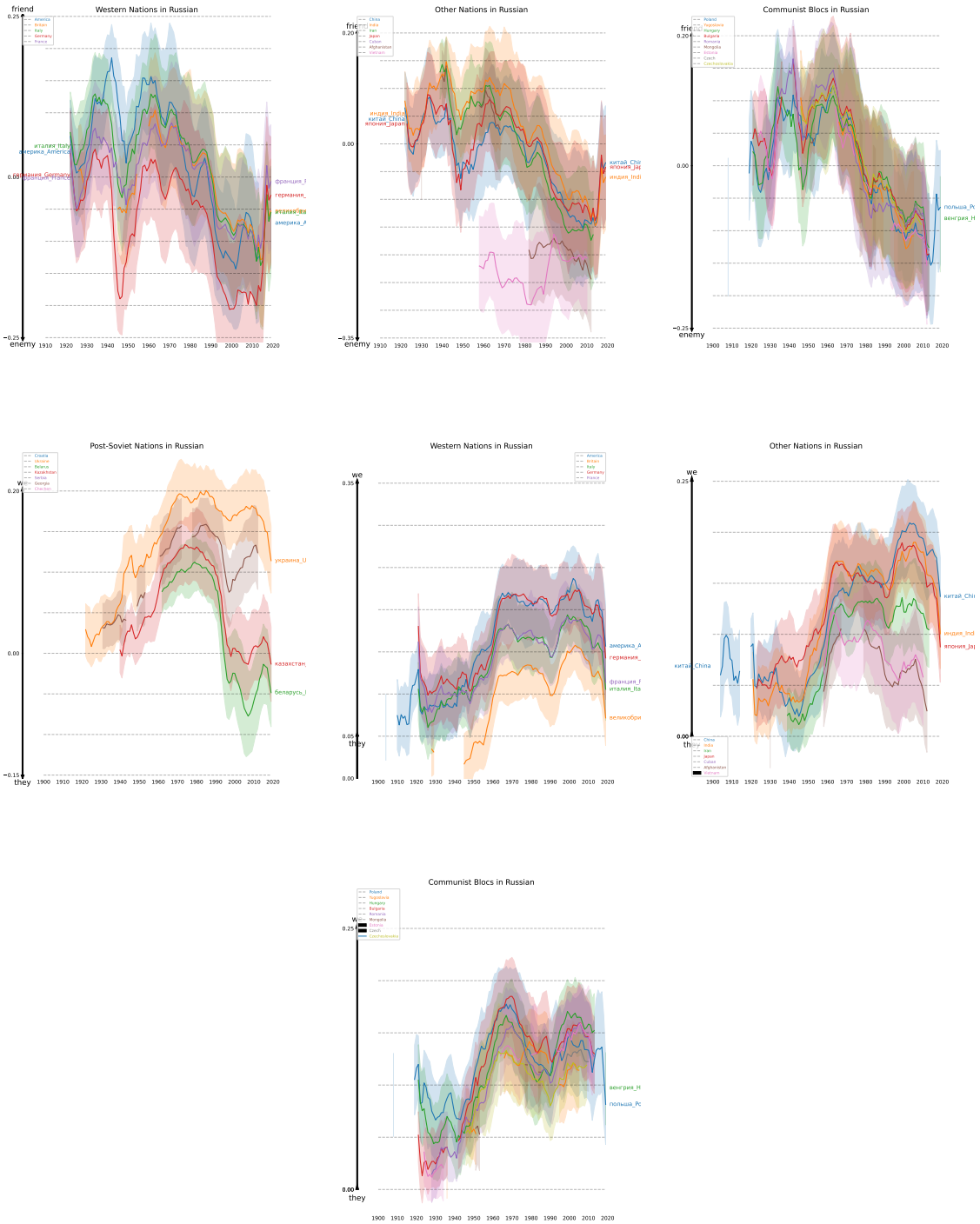

Figure 3.11: Loadings of selected countries in the friend-enemy and we-they dimensions in Russian

stretch of the friend-enemy dimension at some point in history in all other languages. In British English, French, German, and Italian, the stretch all happened during the two world wars. During war time, the word "friend," and "enemy" become more dissimilar to each other in terms of their contextual usage. In Russian, a stretch also occurred, but it seems to correspond to the Great Purge rather than WWII. A stretch in the we-they dimension seemed to occur later in WWII. In other languages, during the same time that the stretch happened, there did not seem to be any significant stretch in the we-they dimensions. US English seems to experience less stretch of dimensions in all time. The distinction of friend and enemy were at their historical low during war time. However, the drop was not as dramatic as it is observed in other languages. It could be due to the fact that the two world wars did not happen on America's home continent and were not existential wars to the United States. Because some spatial distortion seems to also occur in the we-they dimensions of the German, Italian, and Russian spaces, we cannot tell whether the stretch of the friend-enemy dimensions could have anything to do with the spatial distortion observed in German, Italian and Russian plots presented in the preceding section. However, it could still nevertheless have an impact on the dimensional loadings we observed. This could mean that even though we applied the same measure to all years, the measuring ruler itself may not be consistent across years especially during war time. It is like when the special theory of relativity applies to the physical space, not only do things in the space change, but the space itself is changing during those special moments. The same amount of distance in two spaces may not be easily comparable.

Confirmatory analysis

All of our analyses thus far have been based on subjective interpretation. We cannot rule out the possibility that the interpretations only lie in our eyes while in fact they only correspond to some random noises. In the last part of this chapter, we will try to externally

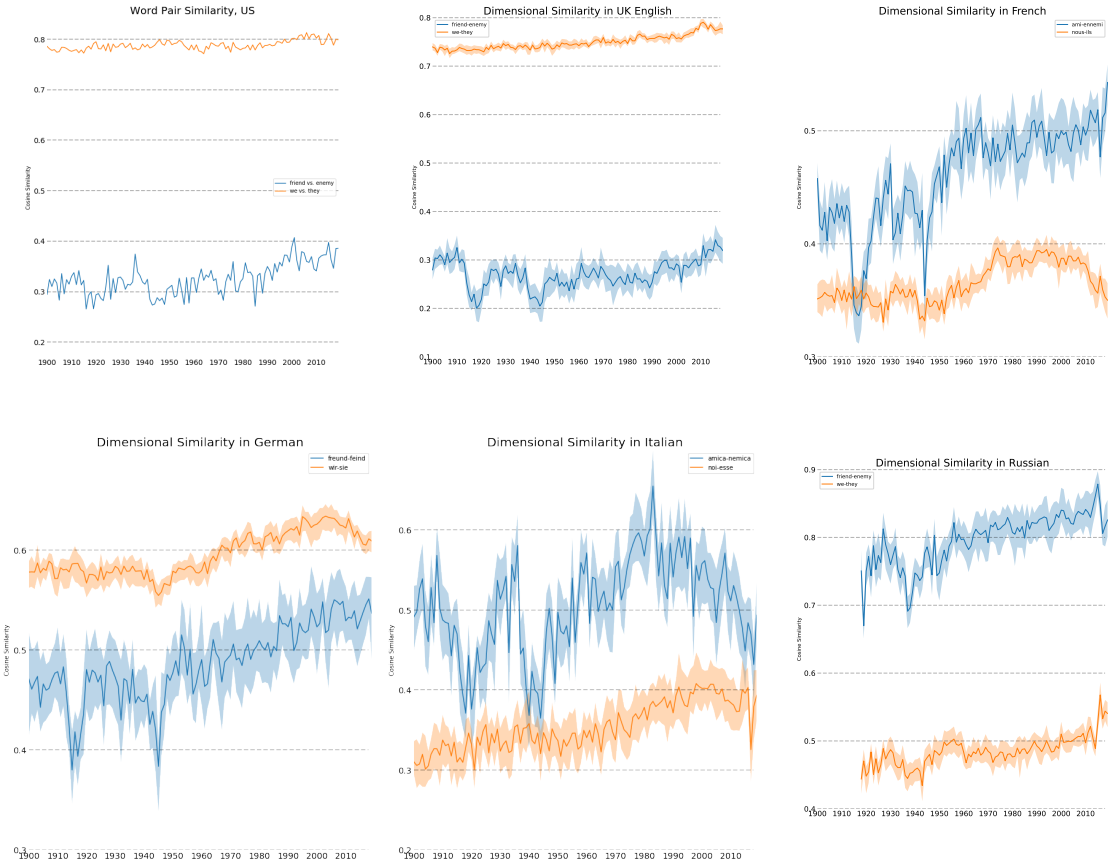

Figure 3.12: Dimensional similarities in 6 languages

validate our measures. In our confirmatory analysis, we run two mixed-effect models on all of the 6 languages to confirm the relationship between perceptions and war and trade. To make our analyses comparable, for each host language, we selected 30 other countries as guest countries for evaluating their friendliness and we-ness. Then, we want to test whether their friendliness or we-ness is correlated with external measures on international trade and war. Although our external dataset almost covers the entirety of the time period we are interested in, many yearly observations are missing due to either lack of statistical records or simply the fact that some nations did not exist in some years. So for each host language, we select 30 guest countries with least missing values. Table 3.3 gives the guest countries being selected into our analysis. For UK and USA, because the abbreviations either do not exist in some languages or were frequently used only in recent time, we used the words "America" and "Britain" to obtain their country loadings. For all other countries, we translated their names using Google Translate.

In our mixed-effect regressions, the units of observations are country-pairs. Because we run separate regressions for each host language, the observations are essentially the guest countries. Based on the graphs presented in the previous section, it looks like in some of the languages, there is some systematic spatial distortion around 2009, which is about the time point Google changed its source of books. To rule out the effect, we only included yearly observations of country-pairs from 1900 to 2009. Missing observations are dropped from the dataset. For host languages with bootstrapped samples, we measured the friendliness and we-ness of each guest country in each sample and pooled all of the observations into a single dataset. For each regression, we included the random effects of all the 30 guest countries, all the observation-years, and all the yearly samples. Even though we controlled for the random effects of every sample, pooling all the samples into a single dataset may still not be the best use of the bootstrapped samples because every guest country is still observed 20 times every year in the pooled dataset. Their errors might still be correlated. In our

Table 3.3: Selected guest countries in our confirmatory analysis

host language	guest countries
US-English	Peru, France, UK, Portugal, Brazil, Indonesia, Chile, Canada, Cuba, Denmark, Italy, China, Argentina, Spain, Japan, Sweden, Colombia, Netherlands, Norway, Egypt, Belgium, Switzerland, Guyana, Finland, Greece, Algeria, Australia, Germany, Philippines
UK-English	USA, Spain, France, Russia, Canada, Indonesia, Sweden, Netherlands, Belgium, Argentina, Chile, Greece, Guyana, China, Finland, Italy, Norway, Ghana, Colombia, Portugal, Brazil, Morocco, Denmark, Algeria, Thailand, Japan, Philippines, Belize, Malta, Egypt
French	Spain, Brazil, Britain, Indonesia, China, Portugal, Canada, Haiti, Italy, Japan, Denmark, Argentina, Cuba, Colombia, America, Egypt, Russia, Chile, Philippines, Peru, Switzerland, Mexico, Belize, Algeria, Netherlands, Norway, Sweden, Greece, Guyana, Belgium
German	USA, Belgium, Netherlands, France, Sweden, Spain, Finland, UK, Denmark, Chile, Portugal, Canada, Argentina, Japan, Norway, Italy, Nigeria, Ghana, Russia, Switzerland, Mexico, Uruguay, Romania, Greece, Algeria, Egypt, Bulgaria, Yugoslavia, Peru, Colombia
Italian	France, Spain, USA, Sweden, Belgium, Netherlands, Chile, UK, Greece, Switzerland, Russia, Portugal, Norway, Argentina, Brazil, Canada, Japan, Germany, Egypt, Finland, Uruguay, Australia, Yugoslavia, Romania, Denmark, Bulgaria, Tunisia, China, Morocco, Indonesia
Russian	Spain, France, UK, Sweden, Belgium, USA, Netherlands, China, Finland, Italy, Greece, Romania, Norway, Portugal, Germany, Egypt, Denmark, Japan, Switzerland, Yugoslavia, Australia, Iran, Brazil, Canada, Bulgaria, Chile, Algeria, Bermuda, Morocco, Tunisia

future analysis, we might consider treating each sample as an imputed dataset in a multiple imputation, running a regression on every sample, and then pooling all the results together with corrected standard errors.

In our regressions, we would like to predict the friendliness and we-ness of guest countries based on war and trade variables. To account for unobserved heterogeneity of guest-countries, we used $\Delta\text{friendliness}_t = \text{friendliness}_t$ - $\text{friendliness}_{t-1}$ as well as $\Delta\text{we-ness}_t$ as our dependent variables. The fixed-effect variables include war_ally (whether a guest country is an ally of the host country in a war), war_enemy (whether it is an enemy of the host country in a war), Δlog_gdp_ratio $= \Delta(\log(\text{gdp}_{\text{host}}) - \log(\text{gdp}_{\text{guest}}))$, Δlog_trade_12, which is the log of the yearly export volume from the host country to the guest country, and Δlog_trade_21, which is the log of the yearly import volume. For both the friend-enemy and we-they dimensions, We included their lagged terms and the other dimensional measure in the fixed effect as well.

Results from US English, UK English and French are reported in Table 3.4 and 3.5. War_enemy consistently has a negative effect on the perceived friendliness of guest countries. The effects are also consistent with our interpretation of the time-series graphs. Being in war with another country significantly reduces the perceived friendliness of that country. War_ally has a positive effect on friendliness in US and UK English. However, the effect sizes are much smaller than that of war_enemy. In French, the ally effect is not present probably due to France's complicated involvement in WWII. We-ness consistently has a positive effect on friendliness, which suggests the two dimensions are somewhat correlated. The effect of trade on friendliness is less clear. In the UK, exporting more leads to less friendliness, and importing more leads to more friendliness. However, in French, the effect of export is in the opposite direction, and the effect of import is null. In the US, neither import nor export has an effect. it is worth noticing that the standard errors are smaller for UK and French simply because there are more observations. Given that the effect sizes are quite weak, it is safer to

conclude that trade has no major impact on friendliness. The same is true for GDP ratio. The significant effect of the lag term suggests that the model may not adequately control for temporal auto-correlation. Including more lag terms might also help us reveal more patterns. The results do not reveal any causal direction although we implicitly assumed that it is more likely that the perceptions are influenced by external events rather than vice versa. On the other hand, Table 3.5 suggests that the we-ness of guest countries is correlated with neither war nor trade.

Regression results from German, Italian and Russian are reported in Table 3.6 and 3.7. Although some effects are significant in some regressions, the signs are not consistent across languages. it is also quite counter-intuitive to interpret some effects. For instance, in Russian, war_enemy leads to more friendliness. As we observed some strange properties in the time series plots, we would rather not further interpret the regression results in those languages.

Conclusion and Discussion

Our investment in constructing perception measures is rewarded with some return. The time series constructed from the US-English, UK-English, and French word-embedding models seem to be highly interpretable. Our interpretations are also consistent with the results of our confirmatory analysis. Some patterns are also highly consistent across these three languages. The friend-enemy measures are highly responsive to wars and perhaps some other major geopolitical events. Some countries, such as Canada and Thailand, are consistently more friendly in the perceptions of these three Western language communities. The host countries' relationships with their former colonies dropped significantly when the colonies successfully declared their independence. There also seems to be some co-variation among all countries in the same region. Because we do not observe any global co-variation, we suspect that these regional co-variations are real, and there could be some high-level patterns beyond the level of nations. For instance, the perceived friendliness of Muslim countries seemed to

Table 3.4: Mixed-effect regression of friendliness in US English, UK English, and French

	Dependent variable:		
	Δfriendliness		
	(US)	(UK)	(French)
war_ally	0.006**	0.003***	0.001
	(0.003)	(0.001)	(0.001)
war_enemy	−0.013**	−0.010***	−0.005***
	(0.006)	(0.001)	(0.001)
$\Delta\text{friendliness}_{t-1}$	−0.456***	−0.482***	−0.479***
	(0.016)	(0.004)	(0.004)
Δwe-ness	0.039**	0.045***	0.045***
	(0.017)	(0.004)	(0.004)
Δlog_gdp_ratio	−0.003	−0.001	−0.002*
	(0.004)	(0.001)	(0.001)
Δlog_trade_12	0.0004	−0.001**	0.001***
	(0.0005)	(0.0003)	(0.0002)
Δlog_trade_21	0.0004	0.0005**	−0.0002
	(0.0005)	(0.0002)	(0.0002)
Constant	−0.00001	0.001	−0.00004
	(0.002)	(0.001)	(0.001)
Observations	2,878	53,563	53,469
Log Likelihood	6,078.754	106,051.200	110,892.300
Akaike Inf. Crit.	-12,133.510	-212,078.500	-221,760.600
Bayesian Inf. Crit.	-12,061.930	-211,971.800	-221,654.000
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3.5: Mixed-effect regression of we-ness in US English, UK English, and French

	Dependent variable:		
		Δwe-ness	
	(US)	(UK)	(French)
war_ally	0.002	−0.002*	−0.003***
	(0.003)	(0.001)	(0.001)
war_enemy	−0.003	0.001	−0.006***
	(0.006)	(0.001)	(0.001)
Δwe-ness$_{t-1}$	−0.519***	−0.483***	−0.481***
	(0.016)	(0.004)	(0.004)
Δfriendliness	0.005	0.042***	0.042***
	(0.016)	(0.004)	(0.003)
Δlog_gdp_ratio	−0.003	0.001	−0.001*
	(0.004)	(0.001)	(0.001)
Δlog_trade_12	−0.0002	−0.0003	−0.0004***
	(0.0005)	(0.0003)	(0.0002)
Δlog_trade_21	−0.00002	0.0003*	0.0004***
	(0.0004)	(0.0002)	(0.0001)
Constant	0.0002	0.001	−0.0001
	(0.002)	(0.001)	(0.001)
Observations	2,878	53,563	53,469
Log Likelihood	6,252.854	108,371.200	115,264.600
Akaike Inf. Crit.	-12,481.710	-216,718.400	-230,505.300
Bayesian Inf. Crit.	-12,410.130	-216,611.700	-230,398.600
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3.6: Mixed effect regression of friendliness in German, Italian and Russian

	Dependent variable:		
	Δfriendliness		
	(German)	(Italian)	(Russian)
war_ally	-0.003^{**}	0.001	−0.003
	(0.001)	(0.001)	(0.012)
war_enemy	−0.0004	-0.010^{***}	0.030^{***}
	(0.001)	(0.002)	(0.005)
$\Delta\text{friendliness}_{t-1}$	-0.496^{***}	-0.493^{***}	-0.475^{***}
	(0.004)	(0.004)	(0.004)
Δwe-ness	0.001	0.023^{***}	0.130^{***}
	(0.004)	(0.004)	(0.008)
Δlog_gdp_ratio	0.001	0.003^{***}	-0.007^{***}
	(0.001)	(0.001)	(0.002)
Δlog_trade_12	−0.00002	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0001)
Δlog_trade_21	0.0002^{**}	0.002^{***}	-0.0003^{**}
	(0.0001)	(0.0003)	(0.0001)
Constant	0.0003	0.0003	−0.001
	(0.001)	(0.002)	(0.004)
Observations	55,251	50,932	39,292
Log Likelihood	108,939.000	99,703.730	61,944.600
Akaike Inf. Crit.	-217,854.100	-199,383.500	-123,865.200
Bayesian Inf. Crit.	-217,747.100	-199,277.400	-123,762.300
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3.7: Mixed effect regression of we-ness in German, Italian, and Russian

	Dependent variable:		
	Δwe-ness		
	(German)	(Italian)	(Russian)
war_ally	-0.002^{*}	-0.003^{**}	-0.016^{**}
	(0.001)	(0.001)	(0.007)
war_enemy	0.002^{**}	0.001	-0.006^{**}
	(0.001)	(0.002)	(0.003)
Δwe-ness$_{t-1}$	-0.498^{***}	-0.494^{***}	-0.487^{***}
	(0.004)	(0.004)	(0.004)
Δfriendliness	−0.004	0.018^{***}	0.037^{***}
	(0.003)	(0.003)	(0.002)
Δlog_gdp_ratio	-0.002^{***}	−0.001	0.001
	(0.001)	(0.001)	(0.001)
Δlog_trade_12	−0.0001	0.0002	−0.0001
	(0.0001)	(0.0002)	(0.0001)
Δlog_trade_21	0.0003^{***}	-0.001^{**}	−0.00002
	(0.0001)	(0.0003)	(0.0001)
Constant	0.0003	0.0002	0.002
	(0.001)	(0.001)	(0.002)
Observations	55,251	50,932	39,292
Log Likelihood	119,841.800	106,389.000	84,735.500
Akaike Inf. Crit.	-239,659.700	-212,754.000	-169,447.000
Bayesian Inf. Crit.	-239,552.700	-212,647.900	-169,344.000
Note:			*p<0.1; **p<0.05; ***p<0.01

all drop at the same time during the 1960s and 1970s. The Iranian Revolution does not seem to be a single coincident that led to the decline. Although we do not have access to the other end of the story, the clash of civilizations between the Western World and the Muslim World seams to have begun way before the end of the Cold War and perhaps even before the First Oil Crisis.

However, results in German, Italian and Russian are much less interpretable. Potentially, many reasons could help explain our failure. First of all, the corpus sizes of these three languages are much smaller. The strange phenomena we observed in their word-embedding spaces could simply be due to randomness in the yearly samples. If the corpus sizes are the cause, fixing the problem wouldn't be too hard. We do have more 5-grams in later years' corpora. For earlier years, we used all we have, but we can pool several year's 5-grams together to build larger corpora. Second, there could be severe selection bias in how the German and Italian corpora were selected. Pro-Fascist books might be banned after the world wars and did not get selected into Google Books. There also existed two Germanys during a long period of time. Authors in East and West Germany might not share the same linguistic dimensions. Results from that period might be misleading. Third, none of the factors listed above could easily explain the spatial distortion observed in the world embedding spaces. They could correspond to some real historical development. However given that we observed similar patterns in English and French corpora before and were able to correct the problem, the problems observed in the German, Italian and Russian spaces are more likely introduced by the word-embedding models. We will still need to understand the inner working of the models better in order to fix the problems.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 1

Table A.1 provides a list of all organizational units that were coded. Organizational units are organized by the domain (ministry, province, party, other) to which they were assigned.

Table A.2 provides statistics for each province. The first column shows the number of times the province was involved in vacancy chains as well as the percentage of all administrative units involved in vacancy chains. The second column shows the number of times the province was involved in isolated transfers as well as the percentage of all administrative units involved in isolated transfers. The third column is the risk ratio of the probability of a vacancy chain hitting a unit over the probability of an isolated transfer hitting the same unit. The fourth column shows the number of times the province appears as a sender in isolated transfers as well as the percentage of times that it appears as the sender (instead of the receiver).

Table A.1: Full list of administrative units by domain

	Administrative Units
Ministry	Administration of the State Material, Aerospace Industry, China Coal Geology Bureau, China National Tourism Administration, Civil Aviation Administration of China, Commission for Science, Technology and Industry for National Defense, Correcting Industrial Illegitimate Practice Office of the State Council, Counsellors' Office of the State Council, Defense Industry Office, Department of Energy, Economic and Trade Office of the State Council, Eighth Ministry of Machinery Industry, Fifth Ministry of Machinery Industry, First Ministry of Machine Industry, Food Safety Commission Office of the State Council, Foreign Affairs Office of the State Council, Foreign Economic Cooperation Department, Foreign Economic and Trade Department, Fourth Ministry of Machine Industry, General Administration of Customs, General Administration of Press and Publication, General Administration of Quality Supervision, Inspection and Quarantine, Hong Kong and Macau Affairs Office of the State Council, Leading Group of Poverty Alleviation and Development Office of the State Council, Legislative Affairs Office of the State Council, Ministry of Agricultural Machinery, Ministry of Agriculture, Ministry of Agriculture and Fisheries, Ministry of Agriculture and Forestry, Ministry of Aviation Industry, Ministry of Aviation and Aerospace Industry, Ministry of Building Materials Industry, Ministry of Business, Ministry of Chemical Industry, Ministry of Civil Affairs, Ministry of Coal Industry, Ministry of Commerce, Ministry of Communications, Ministry of Construction, Ministry of Culture, Ministry of Domestic Trade, Ministry of Education, Ministry of Electronics Industry, Ministry of Environmental Protection, Ministry of Finance / State Revenue, Ministry of Food, Ministry of Foreign Affairs, Ministry of Foreign Economic Relations, Ministry of Foreign Trade, Ministry of Forestry, Ministry of Fuel Chemistry Industry, Ministry of Geology, Ministry of Geology and Mineral Resources, Ministry of Health, Ministry of Housing and Urban-Rural Development, Ministry of Human Resources and Social Security, Ministry of Industry and Information Technology, Ministry of Information Industry, Ministry of Justice, Ministry of Labor, Ministry of Labor and Personnel, Ministry of Labor and Social Security, Ministry of Land Reclamation, Ministry of Land and Resources, Ministry of Light Industry, Ministry of Machinery Industry, Ministry of Machinery and Electronics Industry, Ministry of Materials, Ministry of Metallurgical Industry, Ministry of Nuclear Industry, Ministry of Personnel, Ministry of Petroleum Chemical Industry, Ministry of Petroleum Industry, Ministry of Posts and Telecommunications, Ministry of Power Industry, Ministry of Public Security, Ministry of Railways, Ministry of Science and Technology, Ministry of State Security, Ministry of Textile Industry, Ministry of Transport, Ministry of Urban and Rural Construction and Environmental Protection, Ministry of Water Resources, Ministry of Water Resources and Electricity, Ministry of Weapons Industry, National Audit Office , National Bureau of Statistics, National Committee on Agriculture, National Defense Science and Technology Commission, National Development and Reform Commission, National Exit Inspection and Quarantine Bureau, National Foreign Investment Commission / National Import and Export Management Committee, National Government Offices Administration, National Health and Family Planning Commission, National Infrastructure Construction Commission, National Machinery Industry Committee, National Planning Commission, National Population and Family Planning Commission, National Price Bureau, National Publication Bureau, National Quality and Technical Supervision Bureau, National Radio and Television Administration, National Sports Commission, National Technology Supervision Bureau, National Weather Service, Office of the State Council, Overseas Chinese Affairs Office of the State Council, Production Office of the State Council, Radio and Television Department, Second Ministry of Machinery Industry, Seventh Ministry of Machine Industry, Sixth Ministry of Machine Industry, State (State Council) Administration for Religious Affairs, State Administration for Industry, State Administration for Industry and Commerce, State Administration of Press,Publication,Radio,Film and Television, State Administration of Taxation, State Administration of Work Safety, State Administration of Work Safety, State Board of Education, State Commission for Restructuring, State Council Information Office, State Council Legislative Affairs Bureau, State Council Office for Restructuring the Economic System, State Council Research Office, State Council Taiwan Affairs Office, State Council Three Gorges Project Construction Committee Office, State Development Planning Commission, State Economic Commission, State Economic and Trade Committee, State Environmental Protection Administration, State Environmental Protection Agency, State Ethnic Affairs Commission, State Family Planning Commission, State Food and Drug Administration, State Forestry Administration, State Geological Bureau, State Intellectual Property Office, State Land Administration, State Materials and Equipment Bureau, State Press and Publication Administration, State Science and Technology Commission, State Seismological Bureau, State Sports General Administration, State-owned Assets Supervision and Administration Commission, The Central Bank, Third Ministry of Machine Industry, Water Diversion Project Construction Committee Office of the State Council
Province	Anhui, Beijing, Chongqing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Inner Mongolia, Jiangsu, Jiangxi, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shandong, Shanghai, Shanxi, Sichuan, Tianjin, Tibet, Xinjiang, Yunnan, Zhejiang

Continued on next page

Table A.1: Full list of administrative units by domain

	Administrative Units
Party	CCP Literature Research Center, Central Archives, Central Commission Office for Guiding Cultural and Ethical Progress, Central Financial Work Committee, Central Financial Work Leading Group Office, Central Foreign Affairs Office, Central Organization Department, Central Party History Research, Central Party School, Central Propaganda Department, Communist Youth League, Compilation and Translation Bureau, Guangming Daily, International Department, Ministry of Supervision / Central Discipline Inspection Commission, National Administration for the Protection of State Secrets, People's Daily, Policy Research Office of CPC Central Committee, Political and Legal Affairs Commission, State Commission Office for Public Sector Reform, State Organs Work Committee, The Central Office, United Front, Work Committee of Departments under the CPC CC
Other	All-China Federation of Trade Unions, All-China Women's Federation, China Aerospace Industry Corporation, China Aerospace Science and Technology Group, China Aviation Industry Corporation, China Aviation Industry Group, China Banking Regulatory Commission, China Commercial Aircraft Co., Ltd., China Disabled Persons' Federation, China Earthquake Administration, China Everbright Group, China Federation of Supply and Marketing Cooperatives, China Insurance Regulatory Commission, China National Offshore Oil, China North Industries Group, China Nuclear Industry Group, China Railway, China Red Cross, China Securities Regulatory Commission, China Shipbuilding Industry Corporation, China Shipbuilding Industry Group, China Southern Airlines, China Textile Association, General Office of the CPPCC National Committee, General Office of the NPC Standing Committee, Liaison Office of the Central People's Government in Hong Kong, Liaison Office of the Central People's Government in Macau, National Council for Social Security Fund, National School of Administration, PetroChina, Sinopec, State Council Development Research Center, State Electricity Regulatory Commission, State Grid Corporation, State Power Corporation, Supreme People's Court, Supreme People's Procuratorate, Xinhua News Agency, Xinhua News Agency Hong Kong Branch, Xinhua News Agency Macao Branch

One might argue that the effect we observe is merely the result of long chains containing important positions and short chains containing less important position. The effect could then be the result of people in important positions reaching high ranks. While this does not directly contradict the sponsorship argument, we here still show that even when controlling for the destination of transfers, the effect holds. We do this by switching to the transfer level. Each chained transfer is matched to an isolated transfer that occurs in an eleven-year time window around the chained transfer (5 years to before and after the year in which the chained transfer occurred). Both transfers have the same destination and the people involved have the same party rank at the time of their transfer. Because there are more isolated transfers than chained transfer, we randomly sample matches. We do this 100 times, thus creating 100 samples of matched pairs. The outcome variable of interest is the maximum rank of the person involved in the chained transfer minus the maximum rank of the person involved in the isolated transfer. Figure A.1 shows the histogram of the average rank difference in 100 samples. The differences tend to be positive, demonstrating that individuals involved

Table A.2: Full statistics of the involvement of provincial units in vacancy chains and isolated transfers

	# of Times Involved in Vacancy Chains (% of All Involved Units)	# of Times Involved in Isolated Transfers (% of All Involved Units)	Risk Ratio	# of Times as Senders in Isolated Transfers (% of Times in Isolated Transfers)
Xinjiang	17 (1.6%)	22 (0.6%)	2.52	9 (40.9%)
Chongqing	12 (1.1%)	16 (0.5%)	2.45	7 (43.8%)
Anhui	23 (2.1%)	37 (1.0%)	2.03	22 (59.5%)
Tibet	24 (2.2%)	41 (1.2%)	1.91	20 (48.8%)
Yunnan	19 (1.7%)	33 (0.9%)	1.88	18 (54.5%)
Henan	26 (2.4%)	46 (1.3%)	1.84	24 (52.2%)
Inner Mongolia	18 (1.7%)	33 (0.9%)	1.78	18 (54.5%)
Gansu	22 (2.0%)	41 (1.2%)	1.75	26 (63.4%)
Hunan	18 (1.7%)	34 (1.0%)	1.73	22 (64.7%)
Zhejiang	19 (1.7%)	37 (1.0%)	1.68	26 (70.3%)
Jiangxi	18 (1.7%)	36 (1.0%)	1.63	22 (61.1%)
Guangxi	14 (1.3%)	28 (0.8%)	1.63	9 (32.1%)
Tianjin	15 (1.4%)	33 (0.9%)	1.49	18 (54.5%)
Hebei	26 (2.4%)	59 (1.7%)	1.44	33 (55.9%)
Jiangsu	24 (2.2%)	57 (1.6%)	1.37	37 (64.9%)
Hainan	16 (1.5%)	39 (1.1%)	1.34	7 (17.9%)
Sichuan	24 (2.2%)	61 (1.7%)	1.28	39 (63.9%)
Fujian	17 (1.6%)	44 (1.2%)	1.26	22 (50.0%)
Jilin	18 (1.7%)	47 (1.3%)	1.25	36 (76.6%)
Shanxi	18 (1.7%)	48 (1.3%)	1.23	27 (56.2%)
Qinghai	13 (1.2%)	35 (1.0%)	1.22	21 (60.0%)
Guizhou	15 (1.4%)	41 (1.2%)	1.20	21 (51.2%)
Hubei	19 (1.7%)	54 (1.5%)	1.15	34 (63.0%)
Heilongjiang	19 (1.7%)	55 (1.5%)	1.13	38 (69.1%)
Shaanxi	20 (1.8%)	59 (1.7%)	1.11	28 (47.5%)
Ningxia	11 (1.0%)	33 (0.9%)	1.09	15 (45.5%)
Beijing	28 (2.6%)	88 (2.5%)	1.04	68 (77.3%)
Shandong	19 (1.7%)	63 (1.8%)	0.99	41 (65.1%)
Liaoning	20 (1.8%)	73 (2.0%)	0.90	51 (69.9%)
Shanghai	12 (1.1%)	55 (1.5%)	0.72	35 (63.6%)
Guangdong	11 (1.0%)	70 (2.0%)	0.52	34 (48.6%)

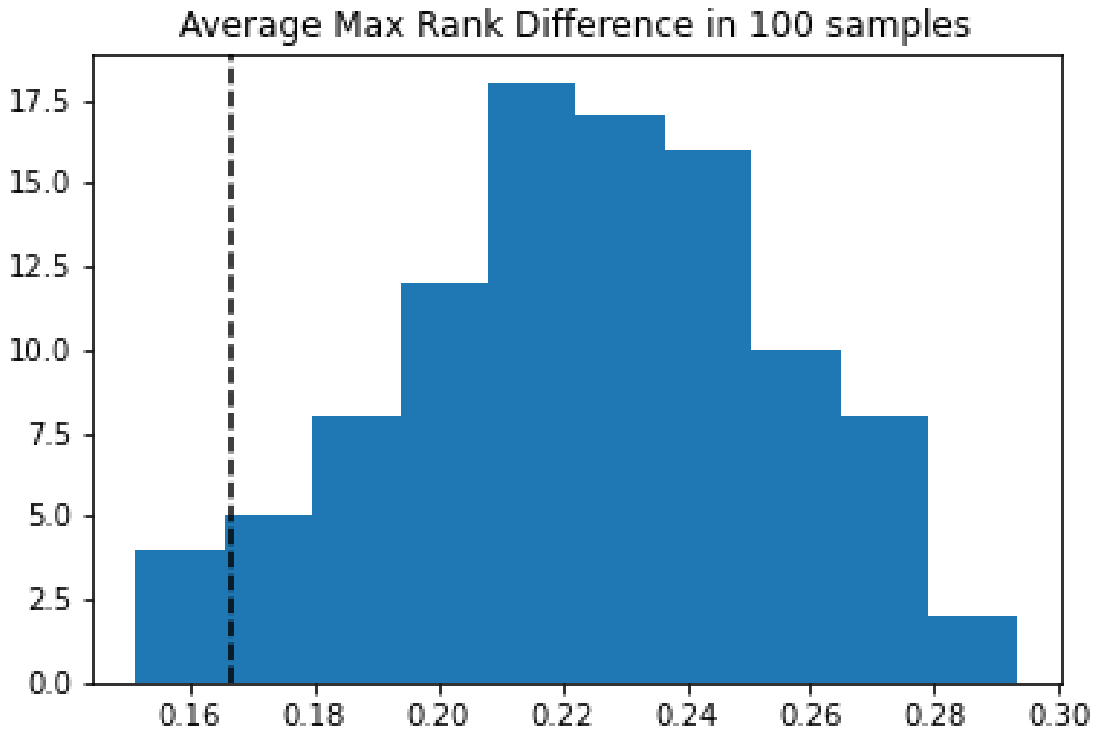

Figure A.1: Robustness check

in chained transfers have systematically higher maximum ranks than their counterparts in isolated transfers.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 2

1. Word Segmentation and Stopword Removal

We initially performed word segmentation following the conventional steps of content analysis. Since our text was written in Chinese, we used the Python module "Jieba", which is an efficient and popular software package for segmenting Chinese text. We used the part-of-speech (pos) segmentation function in the package to cut sentences into word-pos pairs. There are two modes in Jieba: one is the cut-all mode, which cuts sentences into the shortest words, and the other is the default mode, which treats longer word phrases, such as "market economy", as one word. We used the default mode. Then, we removed stopwords using a stopword list created by the Harbin Institute of Technology. We also excluded unwanted words with the following pos tags in Jieba: 'x', 'm', 'p', 'eng', 'c', 's', 't', 'f', 'uj', 'ul', 'u', 'r', 'd', 'uv', 'uz', 'q', 'k' and 'g'. These steps allowed us to eliminate meaningless terms, such as numbers, adverbs and conjunctions (see Table B.1 for more details). The remaining words for all the years were used in the subsequent analyses.

Table B.1: Jieba POS tag table

c: Conjunction	d: Adverbs	f: Locality
g: Morpheme	k: Posterior Component	m: Numeral
p: Prepositional	q: Quantity	r: Pronoun
s: Space	t: Time	uj: Structural Auxiliary
ul: Tense Auxiliary	uv: Structural Auxiliary	uz: Tense Auxiliary
u: Auxiliary	x: Non-morpheme	eng: English words

Note: Jieba POS tags are compatible with the ICTCLAS Chinese POS tagging standard.

2. Correspondence Analysis

Are word frequencies influenced over time by some common factors? To find common trends in word frequencies, we applied correspondence analysis (CA) to the monthly fre-

quency table of all the economic words.

Data Preparation: Tf-idf Weighting

The input matrix $\mathbf{X}$ was a frequency table where x_{ij} was the frequency of the ith word in the jth month. To down-weight words that appeared in many months, and up-weight words that appeared less frequently, we used the term frequency minus the inverse document frequency (tf-idf) weights rather than the raw frequencies as our input. Specifically, the weight of word i in month j was computed as:

$$\text{tf-idf}_{i,j} = \text{tf}_{i,j} \times \text{idf}_i \tag{B.1}$$

$\text{tf}_{i,j}$ is the term frequency of word i in month j, and idf_i is computed as

$$\text{idf}_i = \log \frac{1+n}{1+\text{df}_i} + 1, \tag{B.2}$$

where n was the total number of months included in our analysis, and df_i was the number of months that contained word i (Jones 1972).

Matrix Decomposition

CA, as a variant of principal component analysis, is a classic technique for exploratory analysis of frequency tables. It is especially useful for dimension reduction of categorical and count data with many observational zeros. The goal is, given a count table $\mathbf{X} = [x_{ij}]$, to find row-weight and column-weight vectors $\mathbf{r} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^p$, such that

$$\begin{aligned} r_i &\propto \sum_{j=1}^{p} c_j \frac{x_{ij}}{x_{i\cdot}}, \\ \text{and } c_j &\propto \sum_{i=1}^{n} r_i \frac{x_{ij}}{x_{\cdot j}}, \end{aligned} \tag{B.3}$$

Table B.2: Examples of X − Y ≈ U.S. − Washington D.C (model year=1992)

X	Y	X	Y
France	Paris	Russia	Moscow
U.K.	London	Germany	Bonn
Japan	Tokyo	Australia	Canberra
China	Chinese Government	Egypt	Cairo
India	Delhi	Spain	Madrid

where $x_{i.}$ and $x_{.j}$ are the sums of the ith row and the jth column, respectively. The row scores and column scores in a correspondence analysis are reciprocally determined. The problem can be solved via a generalized singular value decomposition applied to $\mathbf{X}$ (Greenacre 1984).

After a solution is found, the original table can be projected to a low-dimensional space by selecting the left and right singular vectors that correspond to the largest singular values, so that the reduced data preserves maximal variance. The rows and columns can be projected onto one common graph via a biplot, and their distances in the low dimensional space reflect how close they are in the original table $\mathbf{X}$.

3. Vector Algebra in the Word Embedding Spaces

As in the famous "Man - Woman ≈ King - Queen" analogy, vector algebra can be applied to word embedding spaces to learn structural relationships among words (Mikolov, Yih, and Zweig 2013). The following examples demonstrate the effectiveness of our models in preserving contextual relationships in the *People's Daily*. In Figure B.1, selected word pairs are projected from the word embedding spaces to the first two principal components of a PCA. As is shown in the figure, the model is able to capture the analogous relations between provinces in China and their capitals. In Table B.2, the analogous relations between countries and their capitals are shown. Given a country X, we asked the model to find the word Y, such that $X - Y$ was most similar to "U.S. - Washington D.C." in the word embedding space. In most cases, the model was able to return the correct answer.

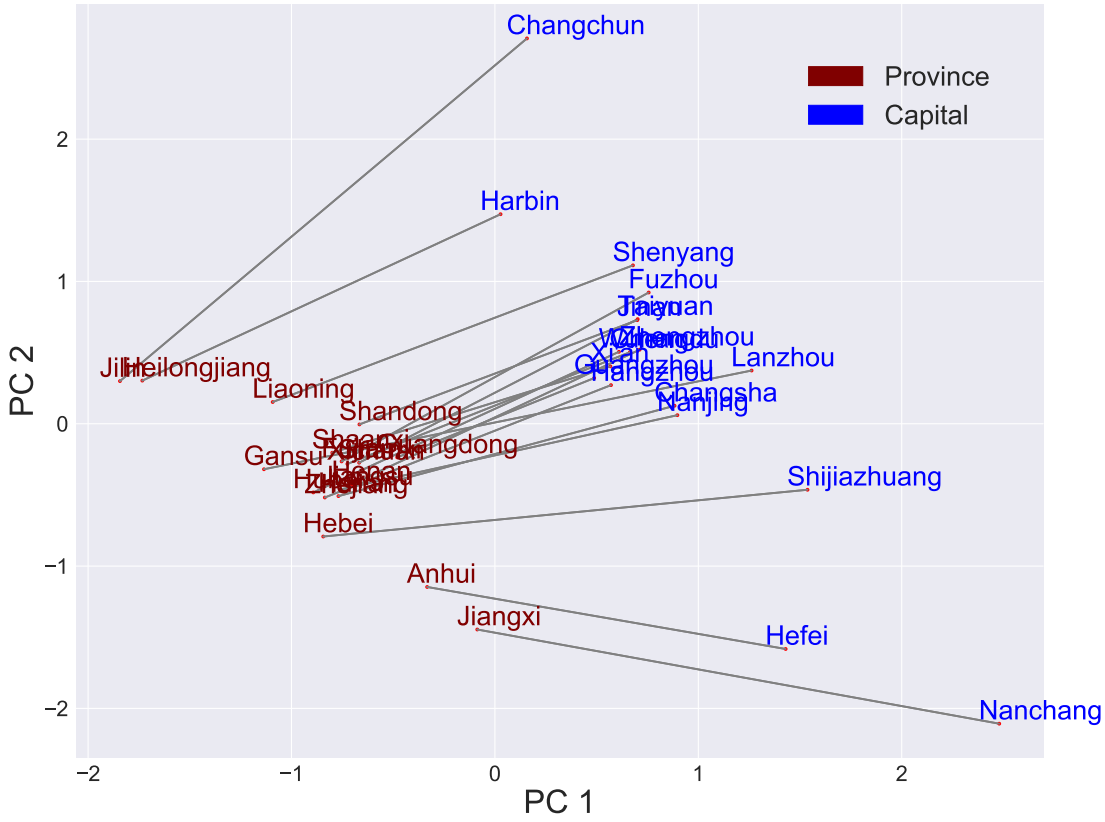

Figure B.1: Two-dimensional PCA projection of the 400-dimensional Skip-gram vectors of provinces and their capital cities in China. As can be seen, the model can automatically capture the similar analogous relations (Year=1992)

4. Procrustes Analysis

Procrustes analysis (PA) was used to align the learned word embedding spaces. Given every year's row-normalized word embedding matrix $\mathbf{W}^{(t)} \in \mathbb{R}^{|\mathcal{V}_t| \times d}$, the rows of which were the input vectors of all the words in that year, the optimal rotation and reflection matrix $\mathbf{R}^{(t)} \in \mathbf{R}^{d \times d}$ was found, such that

$$\mathbf{R}^{(t)} = \underset{\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}}{\arg\min} \left\| \mathbf{W'}^{(t+1)} - \mathbf{W'}^{(t)} \mathbf{Q} \right\|_F \tag{B.4}$$

where $\mathbf{W'}^{(t)}, \mathbf{W'}^{(t+1)} \in \mathbb{R}^{|\mathcal{C}_{t,t+1}| \times d}$ were sub-matrices of $\mathbf{W}^{(t)}$ and $\mathbf{W}^{(t+1)}$, indexed by the common vocabulary set $\mathcal{C}_{t,t+1} = \mathcal{V}_t \cap \mathcal{V}_{t+1}$. After applying a singular value decomposition to $\mathbf{W'}^{(t)\top} \mathbf{W'}^{(t+1)}$, the solution of PA was simply $\mathbf{U}\mathbf{V}^\top$, where $\mathbf{U}$ and $\mathbf{V}$ were the left and right singular vectors (Schönemann 1966). After rotation, word vectors in different years could be aligned to the same vector space.

5. Regression Analysis

Definition of New Word

When we trained our word2vec models, a frequency threshold (= 20) was used to filter out low-frequency words to ensure the qualities of the embeddings. Thus, if a word appeared less than 20 times in a year, it would not appear in that year's word embedding space.

We considered a word to be born in a year t if it appeared in that year's vector space, but had not appeared in the preceding 10 years before t. In other words, the birth year of a word was the first year in which it occurred at least 20 times in the *People's Daily*. If a word stopped occurring for at least 10 consecutive years and then reappeared, we considered it a new word.

Sensitivity Analysis

An alternative measure of a word's instability was $s_t = 1 - \frac{|\mathcal{N}_{t-1} \cap \mathcal{N}_t|}{|\mathcal{N}_{t-1} \cup \mathcal{N}_t|}$, the substitution rate of its k-nearest neighbors, where $\mathcal{N}_t$ was the neighbor set at time tPalla, Barabási, and Vicsek 2007; k was set to be 100 in our study. We re-ran our analyses using this alternative measure. The results are reported in Table B.3. Changing k to 200 did not alter the results. We also compared different choices of k in the original measure in order to ensure the robustness of our results about the field effect. The estimated field effects under the original measure and the alternative measure with $k = 50, 30$ are compared in Table B.4. The effect remained strong no matter which measure was used.

Permutation Test

To further examine how likely the field effect reported in Table 1 of the main article could be a result of randomness, we performed a permutation-based significant test. Specifically, we first randomly re-shuffled the positions of the words in every year's embedding space and calculated their field instabilities given their simulated k-nearest neighbors (k=50). Then, we re-ran Model (2) and obtained an estimate of field effect. The process was repeated 90 times, and the results are reported in Figure B.2. It is clear that the observed field effect is far away from the simulated ones.

6. Field Effect in the First 12 Months

Model

To see whether the field effect was truly exogenous, we regressed all the words' mth month's frequency on field instability for $m = 1, 2, ..., 12$ with the following Poisson regres-

Table B.3: Mixed-effect Poisson regression (1-3) and logistic regression (instability = substitution rate)

Outcome	Frequency			Birth
	(1)	(2)	(3)	
Lagged log frequency	0.382***	0.228***	0.472***	
	(0.004)	(0.005)	(0.006)	
Self-instability		−2.504***		
		(0.007)		
Lagged self-instability			0.586***	
			(0.006)	
Field instability		−2.296***	−3.600***	5.744***
		(0.011)	(0.010)	(0.084)
Initial year × Lagged log frequency	−0.067***	0.019***		
	(0.010)	(0.009)		
Constant	2.065***	6.653***	4.580***	10.702
	(0.035)	(0.037)	(0.030)	(25.850)
Age fixed effects	Yes	Yes	Yes	
Year fixed effects				Yes
Variance of Random Effects				
Word	0.279	0.293	0.170	
Year	0.035	0.039	0.025	
Observations	149,775	149,775	110,613	260,153
Log-likelihood	−2,123,668.000	−1,935,447.000	−1,348,711.000	−123,588.000
Log-likelihood ratio test		376,443.000***		
Bayesian Inf. Crit.	4,247,802.000	3,871,382.000	2,697,875.000	247,638.000

Note: Standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table B.4: Effects of field instability under different instability measures

	Cosine Change		Substitution Rate	
	(k = 50, Original)	(k=30)	(k=50)	(k=30)
Mixed Effect Poisson Regression (Model (2))	−5.600***	−3.334***	−2.296***	−1.048***
	(0.011)	(0.011)	(0.010)	(0.011)
Logistic Regression	6.522***	5.698***	5.744***	5.308***
	(0.092)	(0.084)	(0.084)	(0.078)

Note: Standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

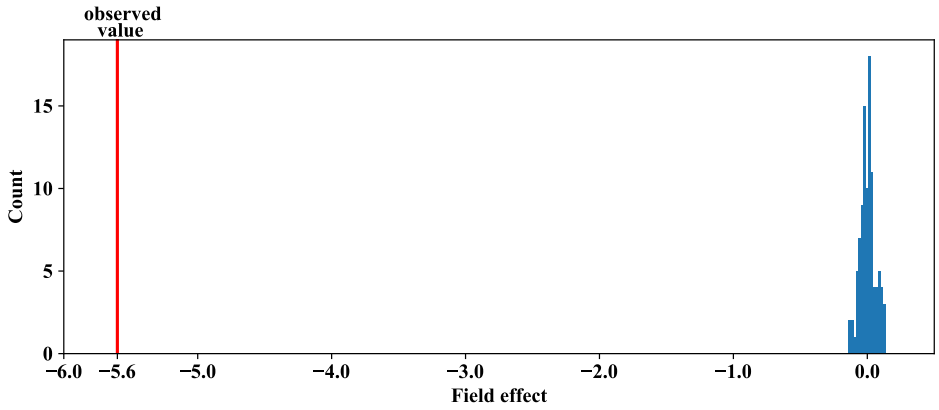

Figure B.2: Histogram of simulated field effects (n=90)

sion:

$$\log \lambda_{i,m} = \alpha_{\text{year}} + \beta x_{i,\text{year}}$$
$$F_{i,m} \sim \text{Poisson}(\lambda_{i,m}), \tag{B.5}$$

where $F_{i,m}$ was the frequency of word i in its mth month, $x_{i,\text{year}}$ was the stability of the field the word i was in at the time, and α_{year} was the intercept for the year.

Sensitivity to Definition of Birth

To test whether our results were sensitive to how birth was defined, we used different case selection criteria to see whether the results differed. Because we used a frequency threshold (=20) when we trained our yearly word2vec models, the first year a word appeared in the word embedding space could differ from the year in which it was truly born. (See also supplemental section 7.) Thus, its observed initial conditions might not be its true initial conditions.

To ensure our results were correct, first, we restricted our analysis to cases in which the first year in the word embedding space was, indeed, the true first year. Only 3412 instances of birth satisfied the strictest criterion. Then, we relaxed the selection criterion to increase the number of cases. The results are shown in Figure B.3. It can be seen that the patterns are similar. No matter which criterion was used, the regression coefficients were positive only

in the first months and then became negative. In other words, the field effect occurred only after birth. It can also be seen that the first-month effect became smaller when more non-first month instances were treated as first-month instances. This observation is consistent with our theory because the field effect had already occurred in those non-first month instances.

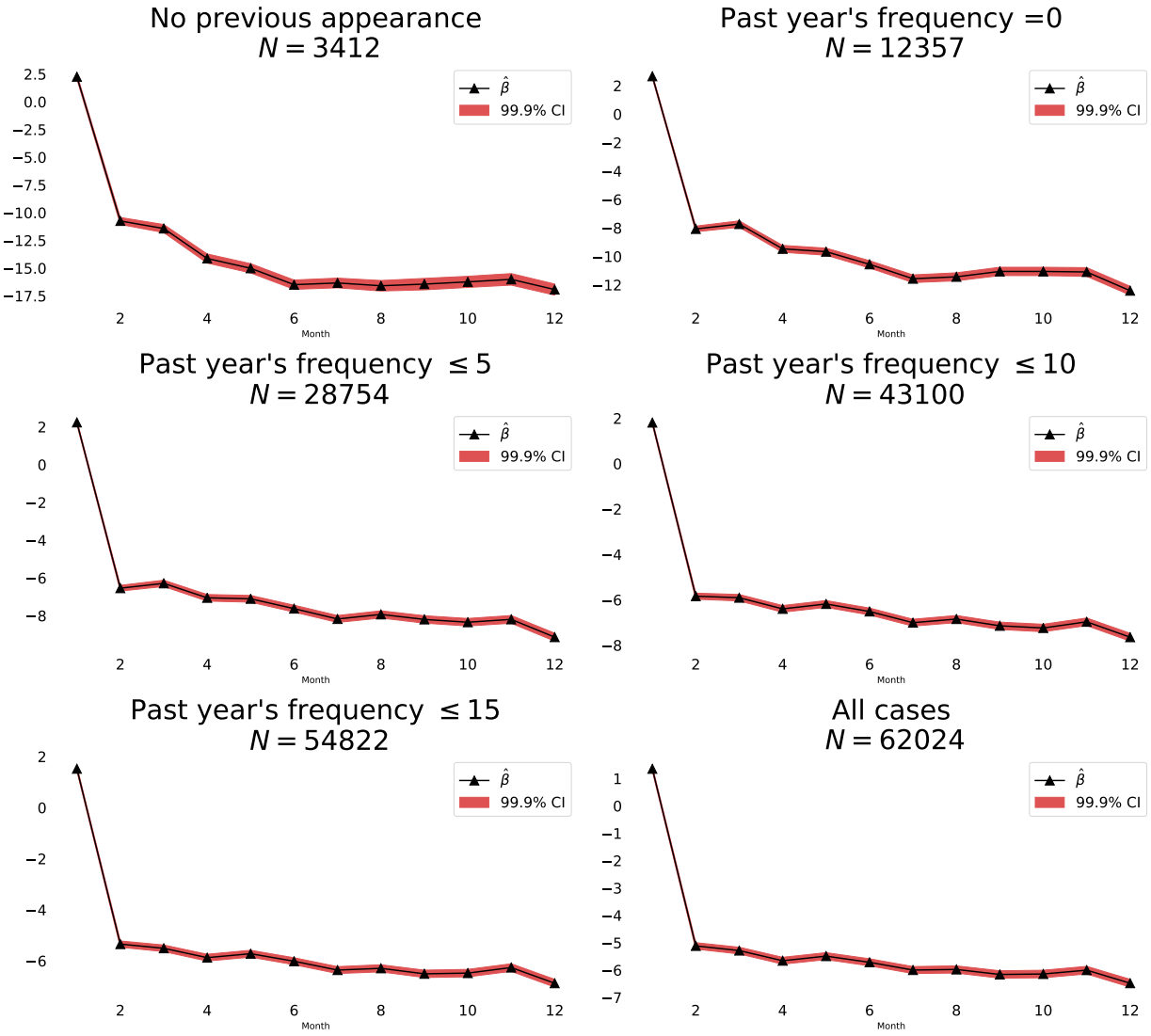

Figure B.3: Effect of field instability on frequency in the first 12 months since inception

Yearly Results

The yearly results are presented in Figure B.4, which shows the bivariate correlations between field instability and log frequency. It can be seen in the figure that the pattern was consistent across all years.

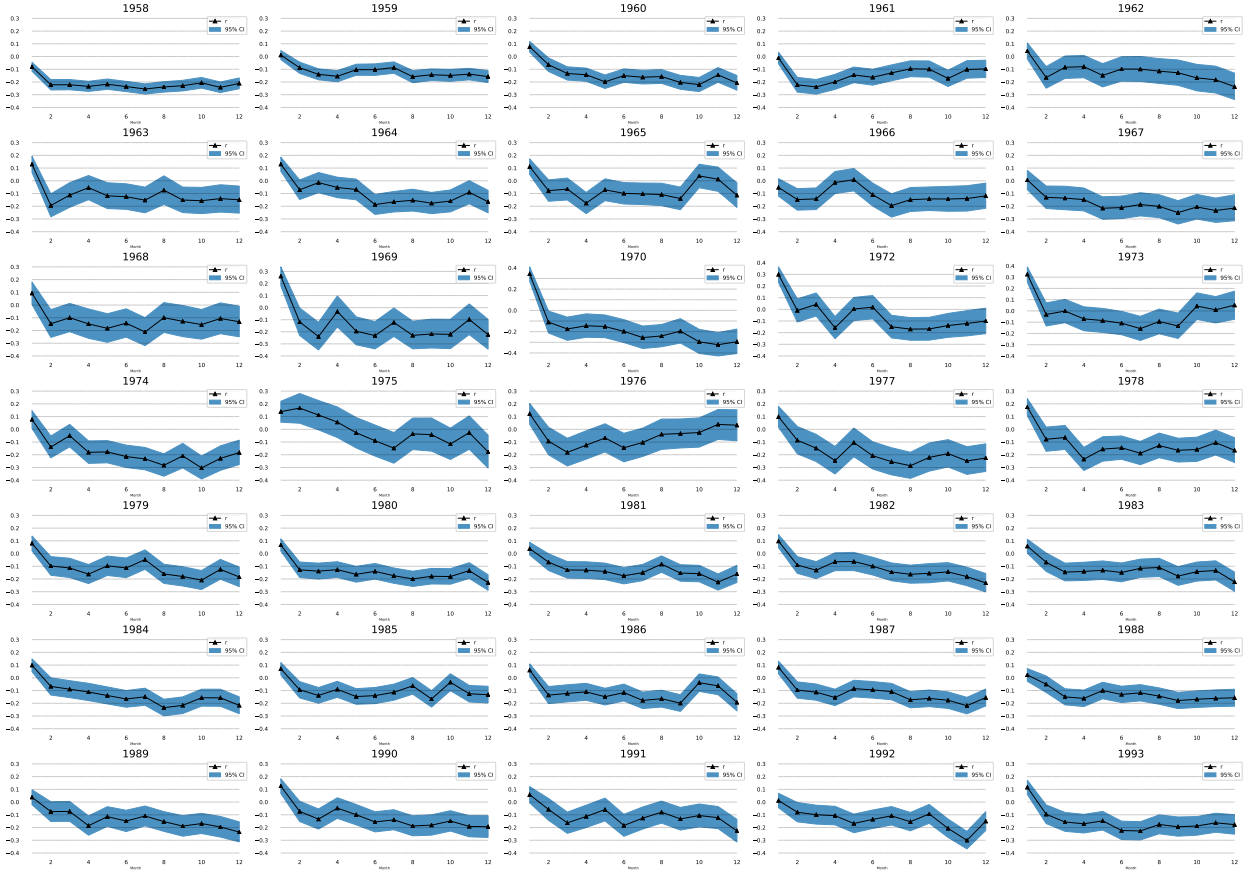

Figure B.4: Correlation between field instability and log frequency in the first 12 months since inception

7. Additional Views of the Vector Space around "Market Economy"

Separate views of the vector space around "market economy" and the estimated instability surface on top of it are given in Figure B.5. The top graph is a two-dimensional t-SNE visualization of the neighborhood of market economy in 1991. The blue arrows trace the

word's positions since its inception. Its neighboring words were plotted in black, their label sizes resemble their frequencies, and the black arrows trace their movements from 1991 to 1992. A polynomial surface of degree 8 was estimated to fit the instability measures of the words in this neighborhood and projected onto the plane. The bottom graph is a 3-dimensional view of the 6th degree surface shown in Figure 5 of the main text.

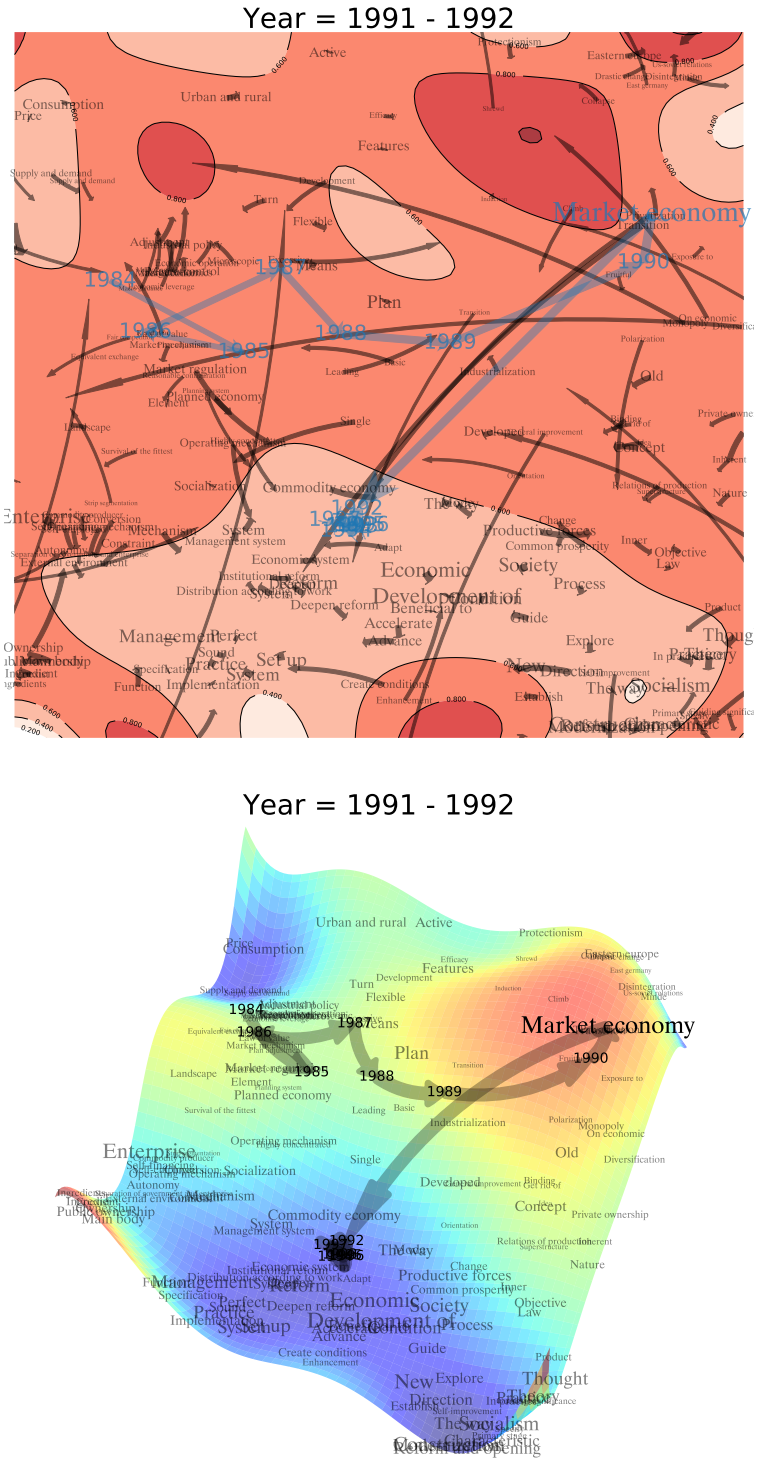

Figure B.5: Separate views of the vector space around “market economy” and the estimated instability surface on top of it

APPENDIX C

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Training Word Embeddings

To train word-embeedings, we used Mikolov, Chen, et al. (2013)'s skip-gram model, which is nowadays the most widely-used neural probabilistic model for representing texts. Given a sequence of training words $w_1, \ldots, w_N$, the objective of the skip-gram model is to find a two-layer representation of words to maximize the average log probability,

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{j=-k}^{k} \log p(w_{i+j}|w_i) \tag{C.1}$$

where k is a window size. The objective function only considers co-occurring word-pairs within a chosen context window. It is in spirit very similar to a semantic network. However, it turns a semantic network into a prediction task. The task is to predict given a word w_i, what other words would appear in its context. Therefor, w_i is considered as the input of the prediction, and w_{i+j} is considered as the output. On paper, it seeks to maximize the log-likelihood by taking the dot product of the vector representations of the input word in the input layer and the output word in the output layer through the softmax function:

$$p(w_i|w_j) = \frac{\exp(\mathbf{u}_{w_i}^{\top} \mathbf{v}_{w_j})}{\sum_{l=1}^{\mathcal{V}} \exp(\mathbf{u}_{w_l}^{\top} \mathbf{v}_{w_j})}, \tag{C.2}$$

where $v_{w_j}, u_{w_i} \in \mathbb{R}^d$ are the input and output vector representations of the word w_j and w_i, and $\mathcal{V}$ is the set of all unique vocabularies in the corpus. The softmax function is essentially the same link function that is used in a multinomial logistic regression. The dot product basically tells the likelihood of each other word appearing in a given word's context window. The higher the dot product is in relation with the dot products of all other pairs, the more likely the pair would appear in the corpus. One can also consider $\mathbf{u}_{w_i}$ as the learned features

of the word w_i and $\mathbf{v}_{w_j}$ as the regression weights for the output word w_j. However, in practice, Mikolov's skip-gram model never actually calculates the softmax function because computing it during training is very costly.

Effect of Corpus Size

As shown in Figure 3.1, the corpus sizes of Google Books are dramatically different over years. In all the Western European languages except for Spanish, years during the Second World War are associated with the least corpus sizes.

It turns out that corpus size also has a huge impact on word-embedding spaces. In my preliminary stage, I trained three rounds of models:

- In the first round, I sampled 1% of the ngrams from every year and trained yearly models.
- In the second round, I used the same corpus as used in round 1 but restricted the vocab size to the same number of most frequent words (n = 11,000) every year. In other words, the model ignored words outside the yearly frequency ranges during training so that the learned vector spaces would have the same number of rows and columns.
- In the third round, I drew a 5% weighted sample from all year's 5-grams. I calculated n_y, the total number of 5-grams in each year y and assigned $1/n_y$ as the weight for all 5-grams in that year. Basically, I over-sampled from smaller years and down-sampled from bigger years so that every year has roughly the same amount of 5-grams drawn into the sample. Then, I trained word-embedding models from the yearly samples.

After each round, I projected some pre-chosen identities to the "friend-enemy" and "we-they" dimensions in each year's word-embedding space using Kozlowski, Taddy, and Evans (2019)'s method. The results are in Figure C.1 and C.2. Initially, by look at the results from round 1, we thought there were some dramatic spatial distortion during war-time. We

were puzzled by the appearance that all countries became more friendly during war time. However, after round 2, some war-time patterns disappear. And after round 3, there is no discernible global war-time patterns that can be observed from the time series. Figure C.3-C.5 are the cosine similarities between "ami (friend)," "ennemi (enemy)," "nous (we)," "ils (they)" as well as "il (he)," "elle (she)," "riches (rich)" and "pauvres (poor)" and all major countries in the world. Each line is the similarity of a country to a chosen word. It seems that all patterns that seemingly happened during war time get eliminated in round 3. When the corpus size is small, word vectors tend to be all closer to each other.

Examining the Qualities of the Embeddings

In computational content analysis, word-embedding is usually used as a unsupervised method for revealing interesting patterns. There hasn't been a clear standard in terms of telling whether a model is good or bad. One way to evaluate how well a model fits the data is to stick with the log likelihood function (which is just the negative of its loss function) and examine whether it is maximized. However, people almost never do it, and the feature is not even properly implemented in the latest version of *gensim*, which is the most widely used package for training Mikolov's skipgram models.

There is a reason why the log likelihood function is somewhat obsolete. The skip-gram models have two implementations for training its word vectors: hierarchical softmax and negative sampling. Neither implementation ever uses the original softmax function. In negative sampling, which is slightly more popular due to its efficiency, the model even completely alters the prediction task from a multi-class problem to a binary classification problem. What it seeks to model is no longer the conditional probability $p(w_{i+j}|w_i)$ but

$$p(y = 1|w_i, w_j) = \text{sigmoid}(\mathbf{u}_{w_i}^\top \mathbf{v}_{w_j}) \tag{C.3}$$

Figure C.1: Dimensional loadings of selected Western nations in the friend-enemy dimension from round 1 (top) to 3 (bottom)

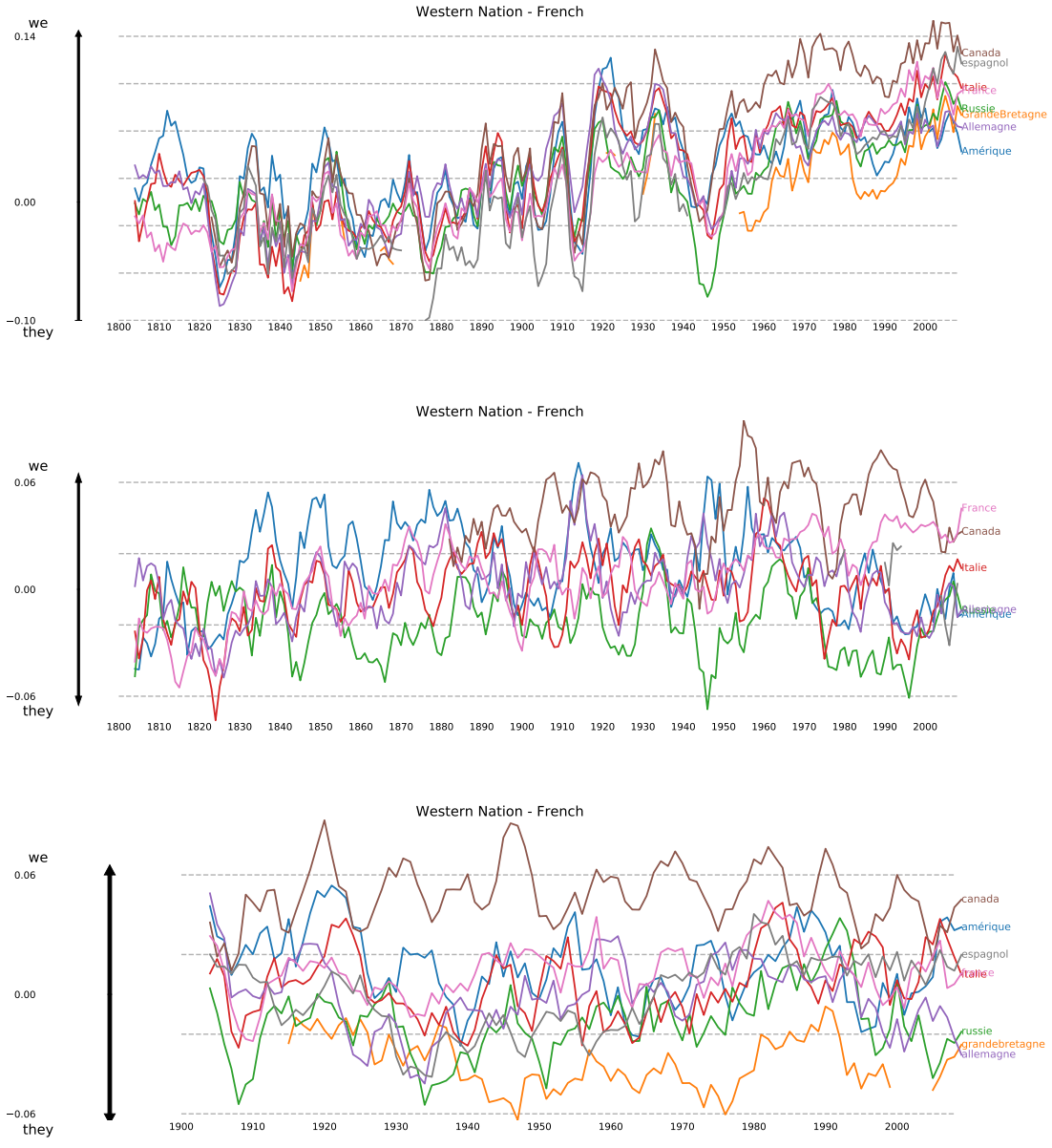

Figure C.2: Dimensional loadings of selected Western nations in the we-they dimension from round 1 (top) to 3 (bottom)

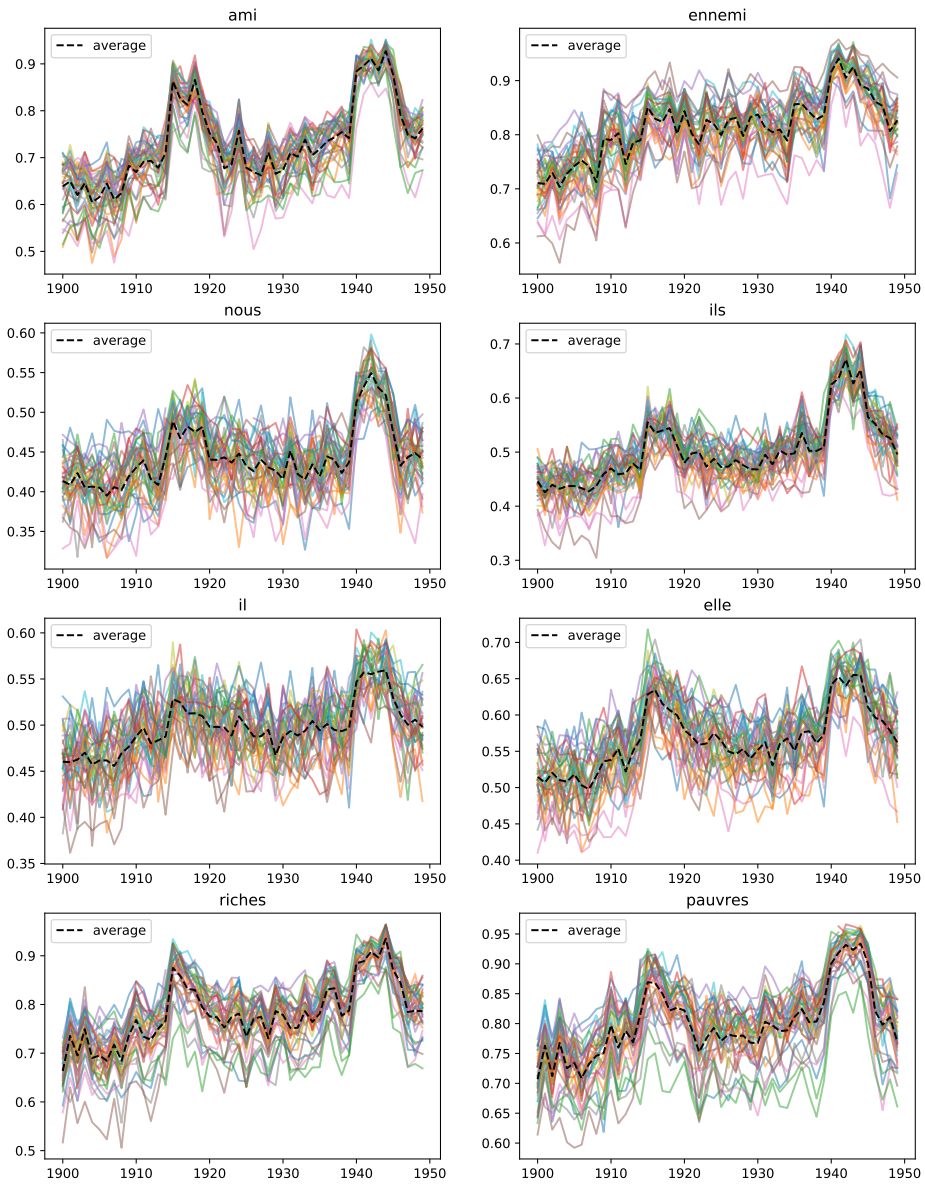

Figure C.3: Country-identity similarities, round 1

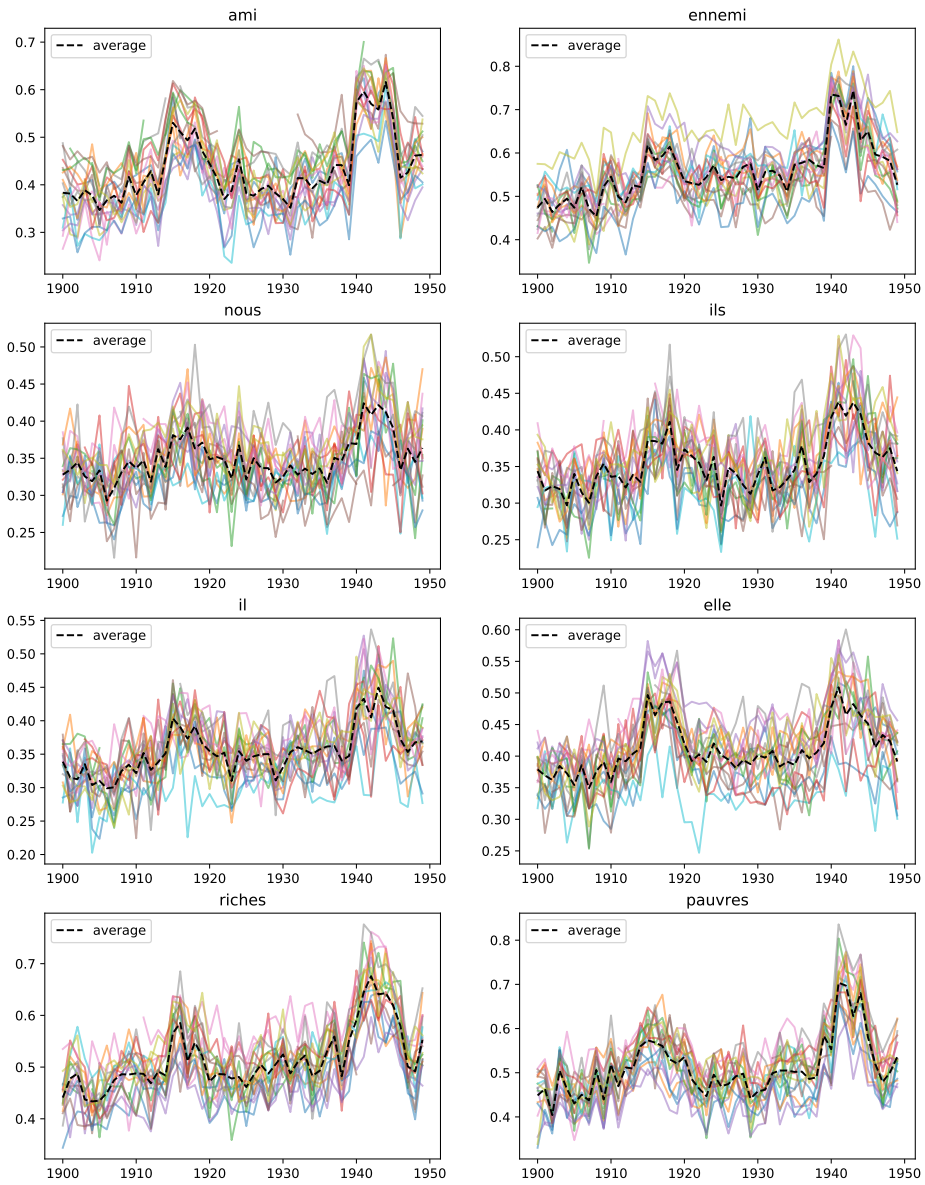

Figure C.4: Country-identity similarities, round 2

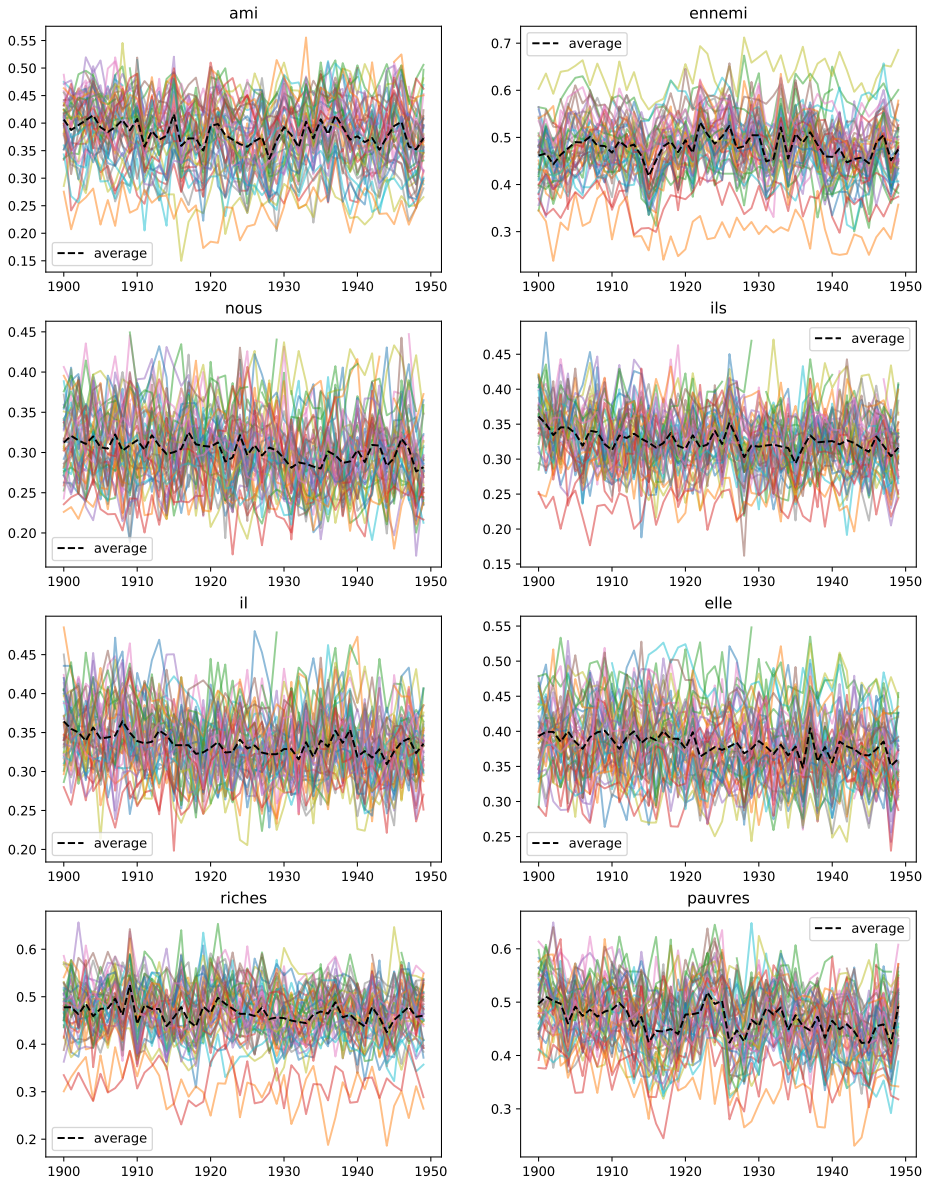

Figure C.5: Country-identity similarities, round 3

, where y is a binary variable indicating whether w_i and w_j co-occur in the training corpus. The sigmoid function is just the link function of a logistic regression. What it does is given a pair of words, take the dot product of the input and output vector representations of the words into a logistic regression to predict whether the pair would actually occur in the training corpus. It has trivial solutions if the length of the input and output vectors are arbitrarily large. So for every positive instance, it adds to the objective function a fixed number of negative instances where the output words are drawn from a noise distribution. Given any pair of words, the logistic regression tells whether the word pair would occur as a true pair in the corpus or simply as a noise pair by chance.

As Mikolov, Sutskever, et al. (2013) explain in their paper, there is no guarantee that vectors trained from negative sampling would maximize the softmax function. However, that is also not the goal of their model. Their goal is to obtain high-quality single-layer vector representation of words. The prediction task is only a means for obtaining word embeddings. One way to simply tell whether a model is good or bad is to use it to obtain similar words that occur in similar contexts, and the results usually intuitively make sense. However, Mikolov also claims that their word embedding models are able to learn some global cultural/linguistic directions such as in the example of "king - queen ≈ man - woman." They found that their models can perform well in solving such analogy tests and constructed a large test corpus (as shown below) including 5 categories of semantic tests and 9 categories of syntactic tests. Their analogy tests have become the standard test for evaluating the quality of word-embedding models. In this study, we also use the Mikolov Analogy Tests to evaluate the performance of our models. We found that negative sampling indeed outperforms hierarchical softmax in analogy tests. Therefore, we use negative sampling for training our models.

capital-common-countries: Athens Greece Baghdad Iraq

capital-world: Abuja Nigeria Accra Ghana

currency: Algeria dinar Angola kwanza

city-in-state: Chicago Illinois Houston Texas

family: boy girl brother sister

gram1-adjective-to-adverb: amazing amazingly apparent apparently

gram2-opposite: acceptable unacceptable aware unaware

gram3-comparative: bad worse big bigger

gram4-superlative: bad worst big biggest

gram5-present-participle: code coding dance dancing

gram6-nationality-adjective: Albania Albanian Argentina Argentinean

gram7-past-tense: dancing danced decreasing decreased

gram8-plural: banana bananas bird birds

gram9-plural-verbs: decrease decreases describe describes

In our experimentation, I came up with two ways to pre-process our ngram data. The original purpose of these tweaks was to save computing resources. In machine learning, there is a general consensus that the more data there is, the more accurate the model trained from the data is. Our results based on analogy tests suggest that adding more variety definitely helps, but reducing repeating information also helps.

Downsampling

The approach I used for speeding up training is downsampling frequent 5-grams. Just like distributions of words, 5-grams also follow a very skewed distribution with a few 5-grams appearing way more times than other 5-grams. Cutting repeated information could help reduce training time. The skipgram model has a built-in downsampling feature that randomly discard high-frequency words during training. Mikolov, Sutskever, et al. (2013) even claim that their downsampling approach could help increase performance in analogy tests in some cases. I tried it. Although it helps to speed up training, it in general worsens performance in analogy tests. I came up with an alternative downsampling scheme. For

5-grams that appear multiple times in a year's corpus, I applied a scaling function to their frequencies such that

$$n'_{ngram_i} = {n_{ngram_i}}^{\beta} \tag{C.4}$$

, where n_{ngram_i} is the count of $ngram_i$ in a yearly sample, and n'_{ngram_i} is the count I used in the training corpus. I varied β from 0 to 1. When $\beta = 1$, there is no scaling. When $\beta = 0$, all counts information is discarded, and every 5-gram is treated as occurring once. When β is between 0 and 1, it heavily downsamples the most frequent 5-gram and lightly downsamples rare 5-grams. The intuition is that in negative sampling, all the model cares is whether a pair of words would co-occur or not in the corpus, repeatedly feeding the model with the same pairs would probably not help with training and could potentially lead to overfitting.

The results of our experiments are in Figure C.6. Overall, there seems to be convincing evidence to suggest that scaling helps. Scale = 0.33 outperforms all other scales in terms of total accuracy. Running the same test on different years of corpus yields similar results. We do not understand why the cubic root is the magic transformation. The breakdown results also suggests that there is no scale that consistently wins in all categories.

Patterns become clearly after all the analogy tests are divided into frequency bins. I calculated the minimum and median of the frequencies of the words in each analogy test. Then I took log base 10 and rounded them into the nearest integers. The bars in Figure C.7 show the average accuracy in each frequency bin. Overall, tests that contain high-frequency words are easier. Scaling also helps in most frequency ranges except for the lowest ones.

Based on the results of the experiments shown in this section, I decide to apply Scale = 0.33 in prepossessing our training data.

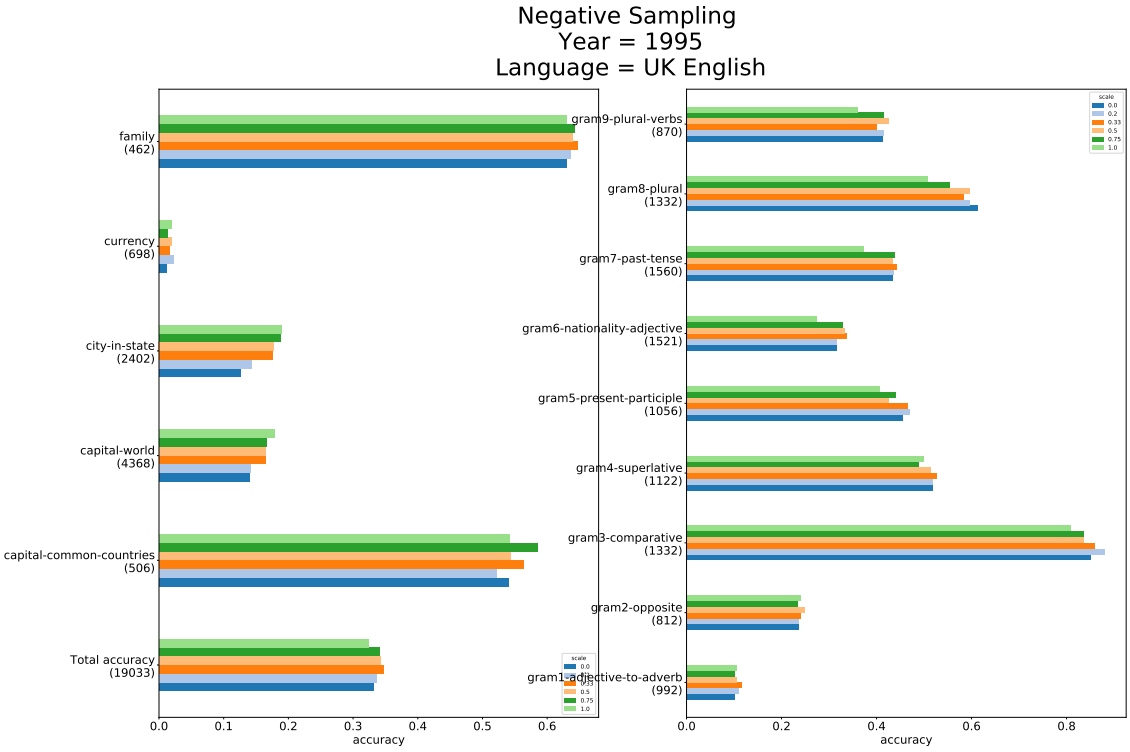

Figure C.6: Breakdown analogy test results in categories

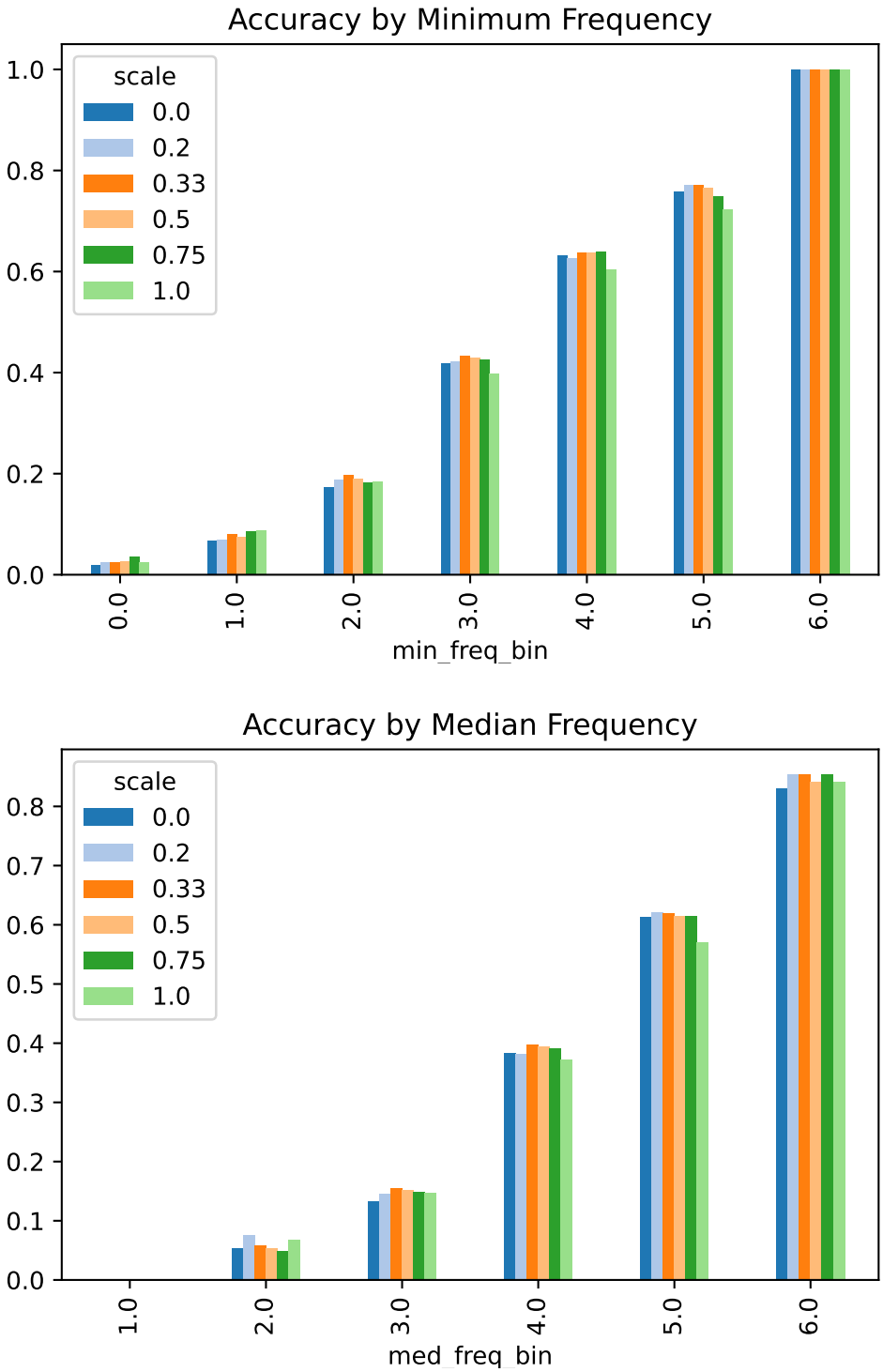

Figure C.7: Breakdown analogy test results in frequency bins

REFERENCES

Abbott, Andrew. 1990. "Vacancy Models for Historical Data." In *Social Mobility and Social Structure,* edited by Ronald L. Breiger, 80–102. Cambridge, UK: Cambridge University Press.

Abbott, Andrew Delano. 2001. *Chaos of Disciplines.* Chicago: University of Chicago Press.

Aberbach, Joel D., Robert D. Putnam, and Bert A. Rockman. 1981. *Bureaucrats and Politicians in Western Democracies.* Cambridge, MA: Harvard University Press.

Adams, Julia. 1996. "Principals and Agents, Colonialists and Company Men: The Decay of Colonial Control in the Dutch East Indies." *American Sociological Review* 61 (1): 467–487.

Alexander, Jeffrey C. 2002. "On the Social Construction of Moral Universals: The 'Holocaust' from War Crime to Trauma Drama." *European Journal of Social Theory* 5 (1): 5–85.

Alexander, Jeffrey C., and Philip Smith. 1993. "The Discourse of American Civil Society: A New Proposal for Cultural Studies." *Theory and Society* 22 (2): 151–207.

———. 2001. "The Strong Program in Cultural Theory: Elements of a Structural Hermeneutics." In *Handbook of Sociological Theory,* edited by Jonathan H. Turner, 135–150. Boston, MA: Springer US.

Amengual, Matthew, and Tim Bartley. 2022. "Global Markets, Corporate Assurances, and the Legitimacy of State Intervention: Perceptions of Distant Labor and Environmental Problems." *American Sociological Review* 87 (3): 383–414.

Bail, Christopher A. 2012. "The Fringe Effect: Civil Society Organizations and the Evolution of Media Discourse about Islam since the September 11th Attacks." *American Sociological Review* 77 (6): 855–879.

———. 2014. "The Cultural Environment: Measuring Culture with Big Data." *Theory and Society* 43 (3): 465–482.

Barkey, Karen. 1994. *Bandits and Bureaucrats: The Ottoman Route to State Centralization.* Ithaca: Cornell University Press.

———. 2008. *Empire of Difference: The Ottomans in Comparative Perspective.* Cambridge, UK: Cambridge University Press.

Bengio, Yoshua, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. 2003. "A Neural Probabilistic Language Model." *Journal of Machine Learning Research* 3 (Feb): 1137–1155.

Berliner, Joseph S. 1957. *Factory and Manager in the USSR.* Cambridge, MA: Harvard University Press.

Blau, Peter M. 1955. *The Dynamics of Bureaucracy: A Study of Interpersonal Relations in Two Government Agencies.* Chicago: University of Chicago Press.

———. 1970. "A Formal Theory of Differentiation in Organizations." *American Sociological Review* 35 (2): 201–218.

Blau, Peter M., and Otis D. Duncan. 1967. *The American Occupational Structure.* New York: Wiley.

Bo, Zhiyue. 2002. *Chinese Provincial Leaders: Economic Performance and Political Mobility since 1949.* Armonk: M.E. Sharpe.

———. 2004. "The Institutionalization of Elite Management in China." In *Holding China Together: Diversity and National Integration,* edited by Naughton Barry J. and Dali L. Yang, 70–100. Cambridge, UK: Cambridge University Press.

———. 2010. *China's Elite Politics: Governance and Democratization.* Singapore: World Scientific.

Bolukbasi, Tolga, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings." *Advances in Neural Information Processing Systems* 29.

Bourdieu, Pierre. 1993. *Sociology in Question.* London: Sage.

Bourdieu, Pierre, and Loic J. D. Wacquant. 1992. *An Invitation to Reflexive Sociology.* Chicago: University of Chicago Press.

Brady, Anne-Marie. 2009. *Marketing Dictatorship: Propaganda and Thought Work in Contemporary China.* Lanham: Rowman & Littlefield Publishers.

Brandt, Loren, and Thomas G. Rawski. 2008. "China's Great Economic Transformation." In *China's Great Economic Transformation,* edited by Loren Brandt and Thomas G. Rawski, 1–26. Cambridge: Cambridge University Press.

Brødsgaard, Kjeld Erik. 2012. "Cadre and Personnel Management in the CPC." *China: An International Journal* 10 (2): 69–83.

Brown, Kerry, and Una Aleksandra Bērziņa-Čerenkova. 2018. "Ideology in the Era of Xi Jinping." *Journal of Chinese Political Science* 23 (3): 323–339.

Burns, John P. 1993. "Administrative Reform in China: Issues and Prospects." *International Journal of Public Administration* 16 (9): 1345–69.

Burns, John P. 1994. "Strengthening Central CCP Control of Leadership Selection: The 1990 Nomenklatura." *The China Quarterly* 138:458–91.

Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition.* Cambridge, MA: Harvard University Press.

Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. "Semantics Derived Automatically from Language Corpora Contain Human-like Biases." *Science* 356 (6334): 183–186.

Callon, Michel. 1998. "The Embeddedness of Economic Markets in Economics." In *The Laws of the Markets,* edited by Michel Callon, 1–57. Oxford: Blackwell.

Callon, Michel, and Fabian Muniesa. 2005. "Peripheral Vision: Economic Markets as Calculative Collective Devices." *Organization Studies* 26 (8): 1229–1250.

Campion, Michael A., Lisa Cheraskin, and Michael J. Stevens. 1994. "Career-Related Antecedents and Outcomes of Job Rotation." *The Academy of Management Journal* 37 (6): 1518–42.

Carlson, Richard O. 1961. "Succession and Performance among School Superintendents." *The Academy of Management Journal* 6 (6): 210–27.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928.* Princeton: Princeton University Press.

Chan, Hon S. 2004. "Cadre Personnel Management in China: The Nomenklatura System, 1990-1998." *The China Quarterly* 179:703–34.

Chandler, Alfred D. 1962. *Strategy and Structure: Chapters in the History of the Industrial Enterprise.* Cambridge, MA: M.I.T. Press.

Chase, Ivan D. 1991. "Vacancy Chains." *Annual Review of Sociology* 17 (1): 133–154.

Chen, Joy, Erik H. Wang, and Xiaoming Zhang. 2021. "Leviathan's Offer: State-Building with Elite Compensation in Early Medieval China," https://doi.org/10.2139/ssrn.3893130.

Chen, Ye, Hongbin Li, and Li-An Zhou. 2005. "Relative Performance Evaluation and the Turnover of Provincial Leaders in China." *Economics Letters* 88 (3): 421–425.

Cohen, Paul A. 1988. "The Post-Mao Reforms in Historical Perspective." *The Journal of Asian Studies* 47 (3): 518–540.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94:S95–120.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent,* edited by David E. Apter, 206–261. New York: Free Press.

Danescu-Niculescu-Mizil, Cristian, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. "Echoes of Power: Language Effects and Power Differences in Social Interaction." In *Proceedings of the 21st International Conference on World Wide Web,* 699–708. WWW '12. Lyon, France: ACM.

DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *American Journal of Sociology* 120 (5): 1473–1511.

Devons, Ely. 1950. *Planning in Practice: Essays in Aircraft Planning in Wartime.* Cambridge, UK: Cambridge University Press.

DiMaggio, Paul. 1997. "Culture and Cognition." *Annual Review of Sociology* 23 (1): 263–287.

DiMaggio, Paul, and Amir Goldberg. 2018. "Searching for Homo Economicus: Variation in Americans' Construals of and Attitudes toward Markets." *European Journal of Sociology* 59 (2): 151–189.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–160. Accessed June 25, 2022.

Dittmer, Lowell. 1978. "Bases of Power in Chinese Politics: A Theory and an Analysis of the Fall of the "Gang of Four"." *World Politics* 31 (1): 26–60.

Dodds, Peter Sheridan, Eric M. Clark, Suma Desu, Morgan R. Frank, Andrew J. Reagan, Jake Ryland Williams, Lewis Mitchell, et al. 2015. "Human Language Reveals a Universal Positivity Bias." *Proceedings of the National Academy of Sciences* 112 (8): 2389–2394.

Doeringer, Peter B., and Michael J. Piore. 1979. *Contested Terrain: The Transformation of the Workaddress in the Twentieth Century.* New York: Basic Books.

———. 1985. *Internal Labor Markets and Manpower Analysis.* Armonk: M.E. Sharpe.

Doyon, Jérôme. 2019. "The Strength of a Weak Organization: The Communist Youth League as a Path to Power in Post-Mao China." *The China Quarterly* 243:780–800.

Duffy, Michael. 1980. "Introduction: The Military Revolution and the State 1500-1800." In *The Military Revolution and the State 1500-1800,* edited by Michael Duffy, 1–9. Exeter: University of Exeter.

Durkheim, Émile. 2001. *The Elementary Forms of Religious Life.* Edited by Mark Sydney Cladis. New York: Oxford University Press.

Eaton, Sarah, and Genia Kostka. 2014. "Authoritarian Environmentalism Undermined? Local Leaders' Time Horizons and Environmental Policy Implementation in China." *The China Quarterly* 218:359–380.

Eccles, Robert G., and Harrison C. White. 1988. "Price and Authority in Inter-Profit Center Transactions." *American Journal of Sociology* 94:S17–S51.

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

Evans, Peter B. 1995. *Embedded Autonomy: States and Industrial Transformation.* Princeton: Princeton University Press.

Fama, Eugene F., and Michael C. Jensen. 1983. "Separation of Ownership and Control." *The Journal of Law and Economics* 26 (2): 301–325.

Firth, J. 1957. "A Synopsis of Linguistic Theory 1930-1955." In *Studies in Linguistic Analysis.* Reprinted in Palmer, F. (ed. 1968) Selected Papers of J. R. Firth, Longman, Harlow. Philological Society, Oxford.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Weixing Wu. 2020. "Social Ties and the Selection of China's Political Elite." *American Economic Review* 110 (6): 1752–1781.

Fiss, Peer C., and Paul M. Hirsch. 2005. "The Discourse of Globalization: Framing and Sensemaking of an Emerging Concept." *American Sociological Review* 70 (1): 29–52.

Fligstein, Neil. 1990. *The Transformation of Corporate Control.* Cambridge, MA: Cambridge University Press.

Foucault, Michel. 1972. *The Archaeology of Knowledge.* New York: Harper & Row.

Fouquin, Michel, Jules Hugot, et al. 2016. *Two Centuries of Bilateral Trade and Gravity Data: 1827-2014.* Technical report. Universidad Javeriana-BogotÃ!‘

Fourcade, Marion, and Kieran Healy. 2007. "Moral Views of Market Society." *Annual Review of Sociology* 33 (1): 285–311.

Freeland, Robert F. 1996. "The Myth of the M-Form? Governance, Consent, and Organizational Change." *American Journal of Sociology* 102 (2): 483–526.

Fukuyama, Francis. 1992. *The End of History and the Last Man.* New York: Free Press.

Gallagher, Michael, and Michael Marsh, eds. 1988. *Candidate Selection in Comparative Perspective: The Secret Garden of Politics.* London: Sage Publications.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes." *Proceedings of the National Academy of Sciences* 115 (16): E3635–E3644.

Geertz, Clifford. 1973. *The Interpretation of Cultures: Selected Essays.* New York: Basic Books.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." *Econometrica* 78 (1): 35–71.

Go, Julian. 2011. *Patterns of Empire: The British and American Empires, 1688 to the Present.* New York: Cambridge University Press.

Goldberg, Amir. 2021. "Associative Diffusion and the Pitfalls of Structural Reductionism." *American Sociological Review* 86 (6): 1205–1210.

Goldberg, Amir, and Sarah K. Stein. 2018. "Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation." *American Sociological Review* 83 (5): 897–932.

Gould, Roger V. 1996. "Patron-Client Ties, State Centralization, and the Whiskey Rebellion." *American Journal of Sociology* 102 (2): 400–429.

Greenacre, Michael J. 1984. *Theory and Applications of Correspondence Analysis.* London: Academic Press.

Gries, Peter, Andrew Fox, Yiming Jing, Matthias Mader, Thomas J. Scotto, and Jason Reifler. 2020. "A New Measure of the 'Democratic Peace': What Country Feeling Thermometer Data Can Teach Us about the Drivers of American and Western European Foreign Policy." *Political Research Exchange* 2 (1): 1716630.

Grusky, Oscar. 1964. "Reply to 'Scapegoating in Baseball'." *American Journal of Sociology* 70 (1): 72–76.

Hamilton, William L., Jure Leskovec, and Dan Jurafsky. 2016a. "Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change." In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing,* 2016:2116. NIH Public Access.

———. 2016b. "Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change." *CoRR* abs/1605.09096.

Harris, Zellig S. 1954. "Distributional Structure." *Word* 10 (2-3): 146–162.

Heclo, Hugh. 1977. *A Government of Strangers: Executive Politics in Washington.* Washington, DC: Brookings Institution.

Ho, Ping-ti. 1962. *The Ladder of Success in Imperial China.* New York: Columbia University Press.

Holbig, Heike. 2013. "Ideology after the End of Ideology. China and the Quest for Autocratic Legitimation." *Democratization* 20 (1): 61–81.

Huang, Yasheng. 2002. "Managing Chinese Bureaucrats: An Institutional Economics Perspective." *Political Studies* 50 (1): 61–79.

Hughes, James M., Nicholas J. Foti, David C. Krakauer, and Daniel N. Rockmore. 2012. "Quantitative Patterns of Stylistic Influence in the Evolution of Literature." *Proceedings of the National Academy of Sciences* 109 (20): 7682–7686.

Iyer, Lakshmi, and Anandi Mani. 2012. "Traveling Agents: Political Change and Bureaucratic Turnover in India." *Review of Economics and Statistics* 94 (3): 723–739.

Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–360.

Jia, Ruixue, and Yiqing Xu. 2018. "Rotating to the Top: How Career Tracks Matter in the Chinese Communist Party," https://doi.org/10.2139/ssrn.3613276.

Jiang, Junyan. 2018. "Making Bureaucracy Work: Patronage Networks, Performance Incentives, and Economic Development in China." *American Journal of Political Science* 62 (4): 982–999.

Jiang, Junyan, Tianyang Xi, and Haojun Xie. 2020. "In the Shadows of Great Men: Leadership Turnovers and Power Dynamics in Autocracies." *Working Paper,* https://doi.org/ https://dx.doi.org/10.2139/ssrn.3586255.

Jones, Karen Spark. 1972. "A Statistical Interpretation of Term Specificity and Its Application in Retrieval." *Journal of Documentation* 28 (1): 11–21.

Kalleberg, Arne L., and Ted Mouw. 2018. "Occupations, Organizations, and Intragenerational Career Mobility." *Annual Review of Sociology* 44 (1): 283–303.

Kane, Anne. 1991. "Cultural Analysis in Historical Sociology: The Analytic and Concrete Forms of the Autonomy of Culture." *Sociological Theory* 9 (1): 53–69.

Kanter, Rosabeth M. 1993. *Men and Women of the Corporation.* New York: Basic Books.

Kaufman, Herbert. 1960. *The Forest Ranger: A Study in Administrative Behavior.* Baltimore: John Hopkins University Press.

Kaufman, Jason. 2004. "Endogenous Explanation in the Sociology of Culture." *Annual Review of Sociology* 30 (1): 335–357.

Kellogg, Katherine C, Melissa A Valentine, and Angéle Christin. 2020. "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals* 14 (1): 366–410.

King, Leslie. 2007. "Charting a Discursive Field: Environmentalists for US Population Stabilization." *Sociological Inquiry* 77 (3): 301–325.

Klimeš, Ondřej, and Maurizio Marinelli. 2018. "Introduction: Ideology, Propaganda, and Political Discourse in the Xi Jinping Era." *Journal of Chinese Political Science* 23 (3): 313–322.

Klingenstein, Sara, Tim Hitchcock, and Simon DeDeo. 2014. "The Civilizing Process in London's Old Bailey." *Proceedings of the National Academy of Sciences* 111 (26): 9419–9424.

Kluver, Alan R. 1996. *Legitimating the Chinese Economic Reforms: A Rhetoric of Myth and Orthodoxy.* Albany, NY: State University of New York Press.

Kou, Chien-wen. 2010. "Paths to the Top: The Political Mobility of Chinese Civilian Leaders after 1987." *Taiwanese Journal of Political Science* 45:1–36.

———. 2014. "The Rise of Youth League Affiliates and Their Paths to the Top." In *Choosing China's Leaders,* edited by Chien-wen Kou and Xiaowei Zang. New York: Routledge.

Kou, Chien-wen, and Wen-Hsuan Tsai. 2014. ""Sprinting with Small Steps" Towards Promotion: Solutions for the Age Dilemma in the CCP Cadre Appointment System." *The China Journal* 71:153–171.

Kozlowski, Austin C., Matt Taddy, and James A. Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review* 84 (5): 905–949.

Kuhn, Philip A. 1980. *Rebellion and Its Enemies in Late Imperial China: Militarization and Social Structure, 1796-1864.* Cambridge, MA: Harvard University Press.

Kulkarni, Vivek, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2014. "Statistically significant detection of linguistic change." *CoRR* abs/1411.3315.

Kullback, S., and R. A. Leibler. 1951. "On Information and Sufficiency." *The Annals of Mathematical Statistics* 22 (1): 79–86.

Lamont, Michèle. 1992. *Money, Morals, and Manners: The Culture of the French and the American Upper-middle Class.* Chicago: University of Chicago Press.

Lan, Zhiyong. 2000. "Understanding China's Administrative Reform." *Public Administration Quarterly* 24 (4): 437–68.

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

Landry, Pierre F., Xiaobo Lü, and Haiyan Duan. 2017. "Does Performance Matter? Evaluating Political Selection Along the Chinese Administrative Ladder." *Comparative Political Studies* 51 (8): 1074–1105.

Lee, Hong Yung. 1991. *From Revolutionary Cadres to Party Technocrats in Socialist China.* Berkeley: University of California Press.

Lee, Suman, and Hyehyun Hong. 2012. "International Public Relations' Influence on Media Coverage and Public Perceptions of Foreign Countries." Public Relations History, *Public Relations Review* 38 (3): 491–493.

Lenin, Vladimir Ilyich. 1929. *What is to be Done? Burning Questions of Our Movement.* New York: International Publishers.

Lévi-Strauss, C. 1983. *Structural Anthropology.* Translated by Monique Layton. Chicago: University of Chicago Press.

Lewis, David E. 2008. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance.* Princeton: Princeton University Press.

———. 2011. "Presidential Appointments and Personnel." *Annual Review of Political Science* 14 (1): 47–66.

Li, Bobai, and Andrew G. Walder. 2001. "Career Advancement as Party Patronage: Sponsored Mobility into the Chinese Administrative Elite, 1949–1996." *American Journal of Sociology* 106 (5): 1371–1408.

Li, Cheng. 2004. "Political Localism Versus Institutional Restraints: Elite Recruitment in the Jiang Era." In *Holding China Together: Diversity and National Integration,* edited by Naughton Barry J. and Dali L. Yang, 29–69. Cambridge, UK: Cambridge University Press.

Li, Cheng, and David Bachman. 1989. "Localism, Elitism, and Immobilism: Elite Formation and Social Change in Post-Mao China." *World Politics* 42 (1): 64–94.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* 89 (9-10): 1743–1762.

Li, Linda Chelan. 2010. "Central-local Relations in the People's Republic of China: Trends, Processes and Impacts for Policy Implementation." *Public Administration and Development* 30 (3): 177–190.

Li, Xiaojun. 2021. "More than Meets the Eye: Understanding Perceptions of China beyond the Favorable–unfavorable Dichotomy." *Studies in Comparative International Development* 56 (1): 68–86.

Lieberson, Stanley. 2000. *Matter of Taste: How Names, Fashions, and Culture Change.* New Haven, CT: Yale University Press.

Lieberthal, Kenneth G. 1992. "Introduction: The 'Fragmented Authoritarianism' Model and Its Limitations." In *Bureaucracy, Politics, and Decision Making in Post-Mao China,* edited by Kenneth G. Lieberthal and David M. Lampton, 1–32. Berkeley: University of California Press.

Lin, Nan. 1999. "Social Networks and Status Attainment." *Annual Review of Sociology* 25 (1): 467–487.

Liu, Yia-Ling. 1992. "Reform from Below: The Private Economy and Local Politics in the Rural Industrialization of Wenzhou." *The China Quarterly* 130:293–316.

Lu, Qinglian. 2018. "Bureaucracy and Networks: The Politics of Career Mobility in Large Organizational Systems." PhD diss., Stanford University.

MacKenzie, Donald A. 2006. *An Engine, not a Camera: How Financial Models Shape Markets.* Cambridge, MA: MIT Press.

Mackenzie, G. Calvin, ed. 1987. *The In-and-Outers: Presidential Appointees and Transient Government in Washington.* Baltimore: John Hopkins University Press.

Manion, Melanie. 1985. "The Cadre Management System, Post-Mao: The Appointment, Promotion, Transfer and Removal of Party and State Leaders." *The China Quarterly* 102:203–233.

Mann, Michael. 2012. *The Sources of Social Power: Volume 1: A History of Power from the Beginning to AD 1760.* New York: Cambridge University Press.

March, James G, and Herbert A Simon. 1958. *Organizations.* New York: John Wiley & Sons.

Martin, John Levi. 2003. "What Is Field Theory?" *American Journal of Sociology* 109 (1): 1–49.

Marullo, Sam. 1985. "Housing Opportunities and Vacancy Chains." *Urban Affairs Quarterly* 20 (3): 364–388.

Marx, Karl. 1974. *The German Ideology.* London: Lawrence & Wishart.

Maskin, Eric, Yingyi Qian, and Chenggang Xu. 2000. "Incentives, Information, and Organizational Form." *Review of Economic Studies* 67 (2): 359–378.

McGregor, Richard. 2010. *The Party: The Secret World of China's Communist Rulers.* New York: Harper.

Mertha, Andrew. 2009. ""Fragmented Authoritarianism 2.0": Political Pluralization in the Chinese Policy Process." *The China Quarterly* 200:995–1012.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83 (2): 340–363.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, null null, Joseph P. Pickett, et al. 2011. "Quantitative Analysis of Culture Using Millions of Digitized Books." *Science* 331 (6014): 176–182.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." *arXiv preprint arXiv:1301.3781.*

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." In *Advances in Neural Information Processing Systems,* 3111–3119.

Mikolov, Tomas, Wen-tau Yih, and Geoffrey Zweig. 2013. "Linguistic Regularities in Continuous Space Word Representations." In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* 746–751.

Moe, Terry M. 1984. "The New Economics of Organization." *American Journal of Political Science* 28 (4): 739.

———. 1985. "The Politicized Presidency." In *The New Direction in American Politics,* edited by John E. Chubb and Paul E. Peterson, 235–71. Washington, D.C.: Brookings Institution.

Mohr, John W, and Brooke Neely. 2009. "Modeling Foucault: Dualities of Power in Institutional Fields." In *Institutions and Ideology,* edited by R.E. Meyer, K. Sahlin, M.J. Ventresca, and P. Walgenbach, 203–255. Emerald Group Publishing Limited.

Mohr, John W. 1998. "Measuring Meaning Structures." *Annual Review of Sociology* 24 (1): 345–370.

Mosca, Gaetano. 1939. *The Ruling Class.* New York: McGraw-Hill.

Mousnier, Roland. 1970. *Peasant Uprisings in Seventeenth-Century France, Russia, and China.* New York: Harper & Row.

Nathan, Andrew J. 2003. "Authoritarian Resilience." *Journal of Democracy* 14 (1): 6–17.

Naughton, Barry. 1995. *Growing out of the Plan: Chinese Economic Reform, 1978-1993.* New York: Cambridge University Press.

Naughton, Barry J, and Dali L Yang. 2004. "Holding China Together: Introduction." In *Holding China Together: Diversity and National Integration in the Post-Deng Era,* edited by Barry J Naughton and Dali L Yang, 1–26. Cambridge, UK: Cambridge University Press.

Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review,* 663–681.

Nee, Victor, and Rebecca Matthews. 1996. "Market Transition and Societal Transformation in Reforming State Socialism." *Annual Review of Sociology* 22 (1): 401–435.

Nelson, Laura K. 2021a. "Cycles of Conflict, A Century of Continuity: The Impact of Persistent Place-based Political Logics on Social movement Strategy." *American Journal of Sociology* 127 (1): 1–59.

———. 2021b. "Leveraging the Alignment between Machine Learning and Intersectionality: Using Word Embeddings to Measure Intersectional Experiences of the Nineteenth Century U.S. South." *Poetics* 88:101539.

Oi, Jean C. 1999. *Rural China Takes Off: Institutional Foundations of Economic Reform.* Berkeley: University of California Press.

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

Ortega, Jaime. 2001. "Job Rotation as a Learning Mechanism." *Management Science* 47 (10): 1361–1370.

Ouyang, Kang. 2001. "Contemporary Development of Marxist Philosophy in China." *Socialism and Democracy* 15 (2): 85–96.

Padgett, John F. 1990. "Mobility as Control: Congressmen through Committees." In *Social Mobility and Social Structure,* edited by Ronald L. Breiger, 27–58. Cambridge, UK: Cambridge University Press.

———. 2012. "The Politics of Communist Economic Reform: Soviet Union and China." In *The Emergence of Organizations and Markets,* edited by John F. Padgett and Walter W. Powell, 271–315. Princeton, NJ: Princeton University Press.

Padgett, John F., and Walter W. Powell. 2012. "From chemical to social networks." In *The Emergence of Organizations and Markets,* edited by John F. Padgett and Walter W. Powell, 92–114. Princeton: Princeton University Press.

Page, Benjamin I., and Robert Y. Shapiro. 1983. "Effects of Public Opinion on Policy." *American Political Science Review* 77 (1): 175–190.

Palla, Gergely, Albert-László Barabási, and Tamás Vicsek. 2007. "Quantifying Social Group Evolution." *Nature* 446 (7136): 664.

Pareto, Vilfredo. 1970. *The Mind and Society.* New York: Harcourt, Brace & Company.

Perrow, Charles. 1961. "The Analysis of Goals in Complex Organizations." *American Sociological Review* 26 (6): 854.

———. 1986. *Complex Organizations: A Critical Essay.* New York: McGraw-Hill Publishing Company.

Perry, Elizabeth J. 1980. *Rebels and Revolutionaries in North China, 1845-1945.* Stanford: Stanford University Press.

Pew Research Center. 2020. "Unfavorable Views of China Reach Historic Highs in Many Countries." https://www.pewresearch.org/global/2020/10/06/unfavorable-views-of-china-reach-historic-highs-in-many-countries/.

Pinfield, Lawrence T. 1995. *The Operation of Internal Labor Markets: Staffing Practices and Vacancy Chains.* New York: Plenum Press.

Poggi, Gianfranco. 1978. *The Development of the Modern State: A Sociological Introduction.* Stanford: Stanford University Press.

Polanyi, Karl. 2001. *The Great Transformation: The Political and Economic Origins of Our Time.* Boston: Beacon Press.

Putnam, Robert D. 1976. *The Comparative Study of Political Elites.* Englewood Cliffs, NJ: Prentice-Hall.

Pye, Lucian W. 1986. "On Chinese Pragmatism in the 1980s." *The China Quarterly,* 207.

Qian, Yingyi, and Barry R. Weingast. 1996. "China's Transition to Markets: Market-preserving Federalism, Chinese Style." *The Journal of Policy Reform* 1 (2): 149–185.

———. 1997. "Federalism as a Commitment to Preserving Market Incentives." *Journal of Economic Perspectives* 11 (4): 83–92.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks: Sage Publications.

Rosenfeld, Rachel A. 1992. "Job Mobility and Career Processes." *Annual Review of Sociology* 18 (1): 39–61.

Rule, Alix, Jean-Philippe Cointet, and Peter S Bearman. 2015. "Lexical Shifts, Substantive Changes, and Sontinuity in State of the Union Discourse, 1790–2014." *Proceedings of the National Academy of Sciences* 112 (35): 10837–10844.

Rumelhart, David E, and Adele A Abrahamson. 1973. "A Model for Analogical Reasoning." *Cognitive Psychology* 5 (1): 1–28.

Saussure, Ferdinand de. 2011. *Course in General Linguistics.* New York: Columbia University Press.

Schmitt, Carl. 1976. *The Concept of the Political.* New Brunswick, NJ: Rutgers University Press.

Schneider, William. 1985. "Peace and Strength: American Public Opinion on National Security." In *The Public and Atlantic Defense,* 321–64. Rowman & Littlefield.

Schönemann, Peter H. 1966. "A Generalized Solution of the Orthogonal Procrustes Problem." *Psychometrika* 31 (1): 1–10. ISSN: 1860-0980. https://doi.org/10.1007/BF02289451.

Shih, Victor, Christopher Adolph, and Mingxing Liu. 2012. "Getting Ahead in the Communist Party: Explaining the Advancement of Central Committee Members in China." *American Political Science Review* 106 (1): 166–187.

Shirk, Susan L. 1993. *The Political Logic of Economic Reform in China.* Berkeley: University of California Press.

Simon, Herbert A. 1947. *Administrative Behaviour.* New York: The Macmillan Company.

Skocpol, Theda. 1985. "Bringing the State Back In: Strategies of Analysis in Current Research." In *Bringing the State Back In,* edited by Peter B. Evans, Dietrich Rueschemeyer, and Theda Skocpol, 3–38. Cambridge, UK: Cambridge University Press.

Smith, D. Randall. 1983. "Mobility in Professional Occupational-Internal Labor Markets: Stratification, Segmentation and Vacancy Chains." *American Sociological Review* 48 (3): 289.

Smith, D. Randall, and Andrew Abbott. 1983. "A Labor Market Perspective on the Mobility of College Football Coaches." *Social Forces* 61 (4): 1147–1167.

Snow, David A. 2004. "Framing Processes, Ideology, and Discursive Fields." Chap. 17 in *The Blackwell Companion to Social Movements,* 380–412. John Wiley & Sons, Ltd.

Sorensen, Aage B. 1974. "A Model for Occupational Careers." *American Journal of Sociology* 80 (1): 44–57.

———. 1977. "The Structure of Inequality and the Process of Attainment." *American Sociological Review* 42 (6): 965.

Spilerman, Seymour. 1977. "Careers, Labor Market Structure, and Socioeconomic Achievement." *American Journal of Sociology* 83 (3): 551–593.

Spillman, Lyn. 1995. "Culture, Social Structures, and Discursive Fields." *Current Perspectives in Social Theory* 15 (1): 129–154.

Steinberg, Marc W. 1999. "The Talk and Back Talk of Collective Action: A Dialogic Analysis of Repertoires of Discourse among Nineteenth-Century English Cotton Spinners." *American Journal of Sociology* 105 (3): 736–780.

Stewart, Philip D., Robert L. Arnett, William T. Ebert, Raymond E McPhail, Terrence L Rich, and Craig E Schopmeyer. 1972. "Political Mobility and the Soviet Political Process: A Partial Test of Two Models." *American Political Science Review* 66 (4): 1269–1290.

Stewman, Shelby. 1975a. "An Application of Job Vacancy Chain Model to a Civil Service Internal Labor Market†." *The Journal of Mathematical Sociology* 4 (1): 37–59.

———. 1975b. "Two Markov Models of Open System Occupational Mobility: Underlying Conceptualizations and Empirical Tests." *American Sociological Review* 40 (3): 298.

Sun, Yan. 1995. *The Chinese Reassessment of Socialism 1976-1992.* Princeton: Princeton University Press.

Swidler, Ann. 1986. "Culture in Action: Symbols and Strategies." *American Sociological Review* 51 (2): 273–286.

———. 2001. *Talk of Love: How Culture Matters.* Chicago: University of Chicago Press.

Taddy, Matt. 2015. "Document Classification by Inversion of Distributed Language Representations." *arXiv preprint arXiv:1504.07295.*

Tajfel, Henri, John C Turner, William G Austin, and Stephen Worchel. 1979. "An Integrative Theory of Intergroup Conflict." *Organizational identity: A reader* 56 (65): 9780203505984–16.

Taylor, M. S., and C. Collins. 2000. "Organizational Recruitment: Enhancing the Intersection of Research and Practice." In *Industrial and Organizational Psychology,* edited by Cary L. Cooper and Edwin A. Locke, 304–34. Oxford, UK: Blackwell.

Tilly, Charles. 1985. "War Making and State Making as Organized Crime." In *Bringing the State Back In,* edited by Peter B. Evans, Dietrich Rueschemeyer, and Theda Skocpol, 169–91. Cambridge, UK: Cambridge University Press.

———. 1992. *Coercion, Capital, and European States, AD 990-1992.* Cambridge, MA: Blackwell.

Treiman, Donald J., and Andrew G. Walder. 2019. "The Impact of Class Labels on Life Chances in China." *American Journal of Sociology* 124 (4): 1125–1163.

Tsou, Tang. 1977. "Mao Tse-tung Thought, the Last Struggle for Succession, and the post-Mao era." *The China Quarterly* 71:498–527.

———. 1983. "Back from the Brink of Revolutionary-"Feudal" Totalitarianism." In *State and Society in Contemporary China,* edited by Victor Nee and David P. Mozingo, 53–88. Ithaca: Cornell University Press.

Turner, Ralph H. 1960. "Sponsored and Contest Mobility and the School System." *American Sociological Review* 25 (6): 855.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* 342 (6157): 468–472.

Vaisey, Stephen. 2009. "Motivation and Justification: A Dual-Process Model of Culture in Action." *American Journal of Sociology* 114 (6): 1675–1715.

van der Maaten, Laurens, and Geoffrey Hinton. 2008. "Visualizing Data Using t-SNE." *Journal of Machine Learning Research* 9:2579–2605.

Van Gunten, Tod. 2019. "Brokers, Clients and Elite Political Networks in Mexico." *SocArXiv* (May). https://doi.org/10.31235/osf.io/vb5gx.

Van Maanen, John. 1984. "Making Rank: Becoming an American Police Sergeant." *Urban Life* 13 (2): 155–176.

Vogel, Ezra F. 2011. *Deng Xiaoping and the Transformation of China.* Cambridge, MA: Harvard University Press.

Walder, Andrew G. 1985. "The Political Dimension of Social Mobility in Communist States: China and the Soviet Union." *Research in Political Sociology* 1:101–17.

Walder, Andrew G. 1995a. "Career Mobility and the Communist Political Order." *American Sociological Review* 60 (3): 309–28.

———. 1995b. "Local Governments as Industrial Firms: An Organizational Analysis of China's Transitional Economy." *American Journal of Sociology* 101 (2): 263–301.

Walder, Andrew G, Bobai Li, and Donald J Treiman. 2000. "Politics and Life Chances in a State Socialist Regime: Dual Career Paths into the Urban Chinese Elite, 1949 to 1996." *American Sociological Review* 65 (2): 191–209.

Walzer, Michael. 1966. *The Revolution of the Saints: A Study in the Origins of Radical Politics.* London, Weidenfeld & Nicolson.

Wang, Xiaonan. 2022. "The Politics of Appointing Insiders and Outsiders: Appointing Provincial Government Agency Heads in China." PhD diss., University of Maryland, College Park.

Weber, Max. 1958. "Politics as a Vocation." In *Max Weber: Essays in Sociology,* edited by Hans H. Gerth and C. Wright Mills, 77–128. New York: Oxford University Press.

———. 1992. *The Protestant Ethic and the Spirit of Capitalism.* New York: Routledge.

White, Harrison C. 1970. *Chains of Opportunity: System Models of Mobility in Organizations.* Cambridge, MA: Harvard University Press.

Wu, Guoguang. 1994. "Command Communication: The Politics of Editorial Formulation in the People's Daily." *The China Quarterly* 137:194–211.

Wuthnow, Robert. 1987. *Meaning and Moral Order: Explorations in Cultural Analysis.* Berkeley: University of California Press.

———. 1989. *Communities of Discourse: Ideology and Social Structure in the Reformation, the Enlightenment, and European Socialism.* Cambridge, MA: Harvard University Press.

Xie, Yu, and Chunni Zhang. 2019. "The Long-term Impact of the Communist Revolution on Social Stratification in Contemporary China." *Proceedings of the National Academy of Sciences* 116 (39): 19392–19397.

Xu, Chenggang. 2011. "The Fundamental Institutions of China's Reforms and Development." *Journal of Economic Literature* 49 (4): 1076–1151.

Yang, Guobin. 2014. "The Return of Ideology and the Future of Chinese Internet Policy." *Critical Studies in Media Communication* 31 (2): 109–113.

Yeung, King-To. 2007. "Suppressing Rebels, Managing Bureaucrats: State-Building During the Taiping Rebellion, 1850–1864." PhD diss., Rutgers University.

Zelizer, Viviana A. 1979. *Morals and Markets: The Development of Life Insurance in the United States.* Columbia University Press.

———. 1994. *Pricing the Priceless Child: The Changing Social Value of Children.* Princeton University Press.

Zeng, Qingjie. 2016. "Control, Discretion and Bargaining: the Politics of Provincial Leader Rotation in China." *Chinese Political Science Review* 1 (4): 623–644.

Zhang, Yang. 2014. "Testing Social Ties against Merits: the Political Career of Provincial Party Chiefs in China, 1990–2007." *Journal of Chinese Political Science* 19 (3): 249–265.

———. 2021. "Why Elites Rebel: Elite Insurrections during the Taiping Civil War in China." *American Journal of Sociology* 127 (1): 60–101.

Zhao, Suisheng. 2016. "The Ideological Campaign in Xi's China: Rebuilding Regime Legitimacy." *Asian Survey* 56 (6): 1168–1193.

Zheng, Yongnian. 2007. *De Facto Federalism in China : Reforms and Dynamics of Central-Local Relations.* Hackensack, NJ: World Scientific.

———. 2010. *The Chinese Communist Party as Organizational Emperor: Culture, Reproduction and Transformation.* London: Routledge.

Zhou, Li-An. 2007. "Governing China's Local Officials: An Analysis of Promotion Tournament Model." *Economic Research Journal* 7:36–50.

Zhou, Xueguang, Yun Ai, Jianhua Ge, Huijun Gu, Ding Li, Lan Li, Qinglian Lu, Wei Zhao, and Ling Zhu. 2021. "The Party–Government Relationship in the Chinese Bureaucracy: Evidence from Patterns of Personnel Flow." *Chinese Journal of Sociology* 7 (3): 315–346.

Zhou, Xueguang, Yun Ai, Jianhua Ge, Huijun Gu, Lan Li, Qinglian Lu, Wei Zhao, and Ling Zhu. 2018. "Stratified Mobility in Chinese Bureaucracy: A Model and Empirical Evidence." *Chinese Journal of Sociology* 38 (3): 1–45.