

New Wine in Old Bottles: Ideological Transformation
and the Rhetorical Creation of the Market in China's
People's Daily

Shilin Jia¹ & Linzhuo Li¹

June 1, 2021

¹Department of Sociology, University of Chicago, Chicago, IL 60637, USA

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

— Hegel, *The Phenomenology of Spirit*

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 the 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.

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 (Brandt and Rawski 2008; Shirk 1993). 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. Numerous studies have explored, in detail, the political and economic conditions of this transformation (Shirk 1993; Naughton 1995; Nee 1989; Qian and Weingast 1996; Walder 1995; Oi 1999).

Yet, how exactly could this transformation have happened on the rhetorical level? How could a new ideological regime supersede an old one? How could a belief system be reoriented and incorporate ideas that it denied in the past?

Many existing explanations of China’s economic transformation implicitly assume that the questions we just raised don’t matter. Political and economic reforms started first. Ideological transformations then followed. Some scholars insist that there is an internalist story to be told (Tsou 1977, 1983; Sun 1995; Kluver 1996). However, as humans have a natural tendency to rationalize what they see, it is hard to say whether any internalist interpretation is a product of confirmation bias. In sociology, there has been a long debate about whether ideology and culture have any independent explanatory power regarding historical processes (Kaufman 2004) The theoretical controversy can be best captured in sociological analyses of the Protestant Reformation. China’s redefinition of socialism, in some way, is the Reformation of the 20th century. Whether the discourse change can be internally explained is an old problem in its new realization. In this essay, we propose an empirical framework and a specific research design to adjudicate the issue. 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-probablistic models, we constructed the rhetorical fields in which ideas were embedded in and provided an internalist explanation of how rhetorical change was conditioned by the structures of the fields.

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 field. Through confirmatory analysis, we demonstrate that 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. Regardless of political turmoil, the Party’s state 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. Ideology presupposes a high degree of rationalization of ideas, which might be lacking among ordinary people. However, Ideological constraints can be particularly strong among a society’s elite members in demand for coherent and consistent understandings of the world (Converse 1964).

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 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 soviet-style social planning (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” is solely a capitalist concept and has no place at all in a socialist economy. The CCP’s later embrace of it is apparently a huge reversal in its official doctrinaire. 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 (Nee 1989). 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). Fiscal decentralization and promotional incentives induced local governments to act as *de facto* enterprises in an industrial competition (Walder 1995; Oi 1999; Li-An 2007; Qian and Weingast 1996, 1997). 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 (Sun 1995). The rhetorical transformation took place through stages that depended on the conclusions of the previous stages.

According to this line of research, the CCP’s discursive change began with a negation period that reverted extreme Maoist leftism to pre-1956 socialism. During the first stage, the most important ideological accomplishment was establishing “practice as the sole criterion of truth” through a major debate in party history. Only after establishing the new epistemological ground, the reform entered into the next stages of debating what constitutes the core elements of socialism. Legitimation of people’s self-interests and de-emphasis of public ownership only took place step by step. The development of socialism was gradually rephrased and equated to the “scientific development of productive force.” Then, by situating China in an “early stage of socialism,” party theorists started to argue that state planning is not an essential element of socialism. The concept of “market economy” as a socialist means of distribution was finally legitimized in 1992. The transformation can hardly be imagined without the earlier stages of theoretical development. Scholars advocating for the independent role of ideology dismiss the theses that ideology is purely a means of political struggle and political elites can easily change their ideological commitment at will. They observe that the CCP’s rhetorical change happened incrementally through zigzags, reject the simplification that the Chinese transformation was a new economic model superimposing an old one, and see a continuation in this process.

Such internalist explanations sound plausible. 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). Can subjective interpretations have any objective ground?

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 sociological discussions of ideological change, because of the special role the Protestant Reformation played in Max Weber’s sociological inquiry and founding of the discipline, the Reformation has a special role of garnering nearly all important debates about the role of ideology in the development of history. We believe that none of the empirical questions we pose in this essay is new. By answering our questions about the Chinese ideological transformation, We can also help address many old problems. We would therefore like to take a detour first and review some of the most important sociological debates about ideological transformation.

Protestant Reformation Revisited

From a Marxist perspective, an ideological change takes the form of a 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 desires to demonstrate that

they were elected by God. Calvinism was conducive to the early development of Capitalism because there was an “elective affinity” among those ideas.

The historical accuracy of Weber’s account remains to be a subject of debate. Weber might have been cherry-picking evidence (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 1948). 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. When one social sphere is in conflict with another social sphere, 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 for escaping the tension, and therefore the ideological change was predictable and followed an “inner law” of its own.

In “Culture in Action,” 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.

Swidler makes a distinction between “settled lives” and “unsettled lives.” During settled periods when there is no structural need for social change, culture is less coherent and provides ritual traditions that regulate actions. During unsettled periods when great social transformations are ongoing, different ideologies emerge in competition with each other and articulate in more consistent fashions. Whether one ideology thrives depends on whether it

serves a structural need well.

A closer read of Walzer (1966)'s book *The Revolution of the Saints* nevertheless suggests that there was more beyond a structural need. Walzer dismisses the Webertain argument that there was a historical link between the 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.

Robert Wuthnow (1989) theorizes 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. In *Communities of Discourse*, Wuthnow gives a detailed analysis of the social context in which the Reformation arose. In short, Europe's commercial expansion over the sea gave the rise of 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, Wuthnow's book provides much fewer empirical details and evidence about the field story than the social-structuralist story.

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 propose an empirical framework to understand the second question.

How to Endogenously Explain Culture?

Among cultural sociologists, although there exist 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 (1948) 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 towards treating meaning too seriously (Geertz 1973, 195). For instance, although social contexts and discursive fields allegedly play equal weight 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 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 constructs. In sociology, field as a concept has already been deployed in internalist explanations of culture (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*, 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 of 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 occupies the position.

In particular, we propose that a new word has a better chance to survive 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 1948). 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,

Hypothesis 1 *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 the test of 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 "watch dog," 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 were more interested in changes in substantive content rather than linguistic styles; thus, only non-functional words were used in our analyses (See Supplementary Section 1). After the frequencies of all the words of interest were counted, a probability distribution of the words was derived as a bag-of-words representation of the corpus. Temporal changes were our primary focus. The probability distribution of words between two time periods was compared by calculating the Kullback-Leibler (KL) divergence. Furthermore, we employed 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 allowed 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 were ignored in our bag-of-words models, we employed 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 method 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). The 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 provided a way for us to select domain-specific keywords. More specifically, we used the nearest neighbors of the words “economy” and “politics”, respectively, as economic keywords and political keywords. These domain-specific words allowed 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; William L. 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 took its k-nearest neighboring words in the vector space in year t . Then, we computed the instabilities of the neighboring words from $t - 1$ to t and took their average as the instability of the field. 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 1 gives a example of a stable field and an unstable field under each measure. Stable fields are neighborhoods where there are less movements and less new neighbors. Conversely, unstable fields are neighborhoods with more movements and more new neighbors. Our measures should make an intuitive sense.

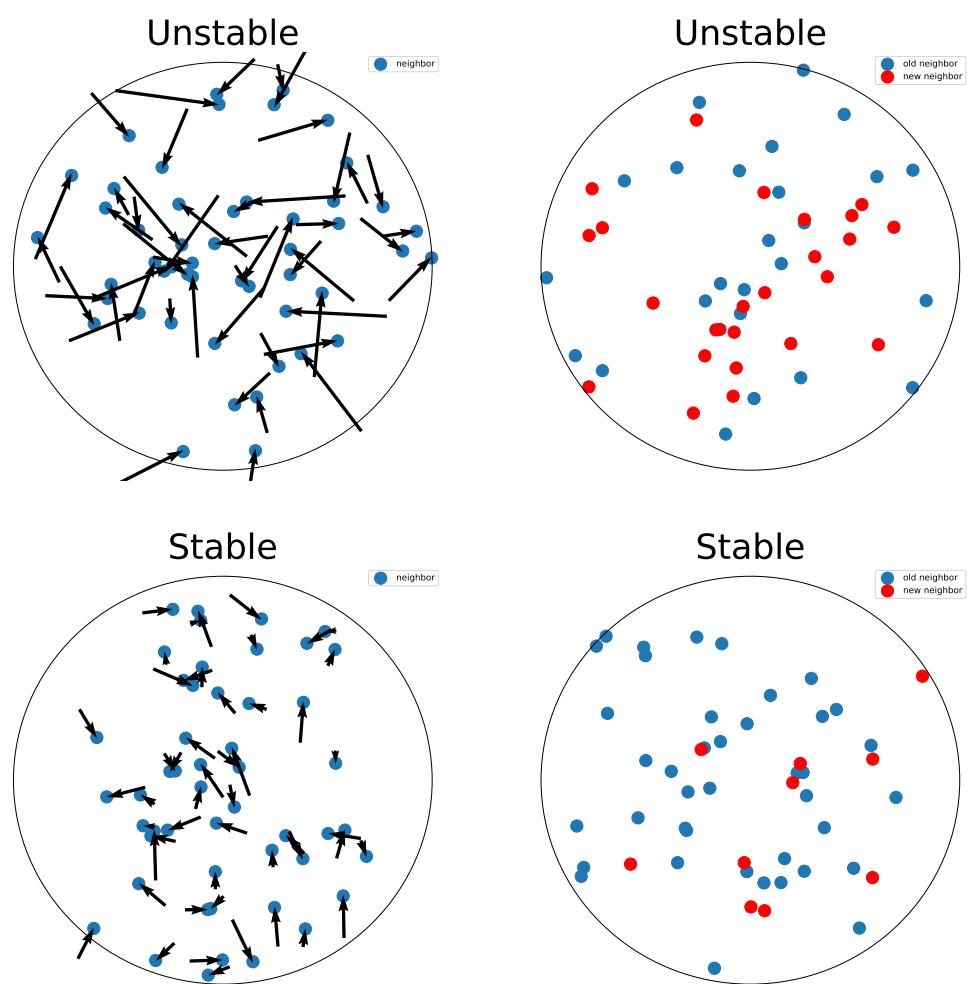


Figure 1: 2D examples of stable and unstable fields under two measures of field stability

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 CCP’s state ideology, especially in the economic domain. Heatmaps, based on year-to-year KL divergences, are shown in Figure 2. The analysis was 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.

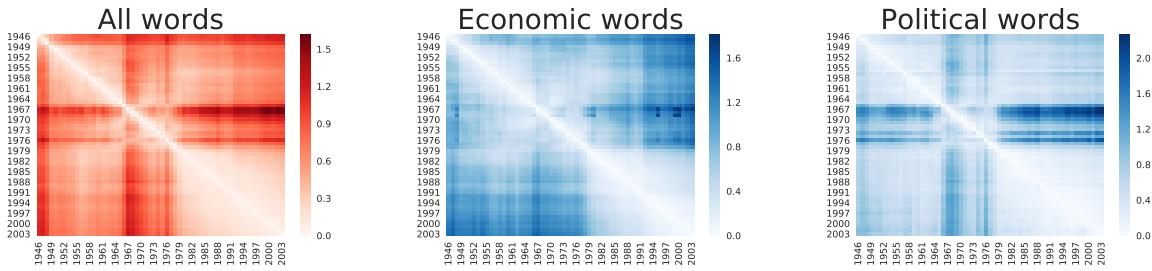


Figure 2: Heatmaps based on year-to-year KL divergences

Time series, which were constructed with two choices of prior distribution, are plotted in Figure 3. The time series on the top of the figure were 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 were 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 had a much shorter memory during the reform period than the previous period. After an external shock, it appeared to quickly revert back to its mean. For instance, in Figure 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 had very little long-term impact on the CCP’s overall state discourse. In the economic domain, the shock was hardly detectable. Quickly after the crackdown, the CCP’s official state rhetoric proceeded as if Tiananmen never happened.

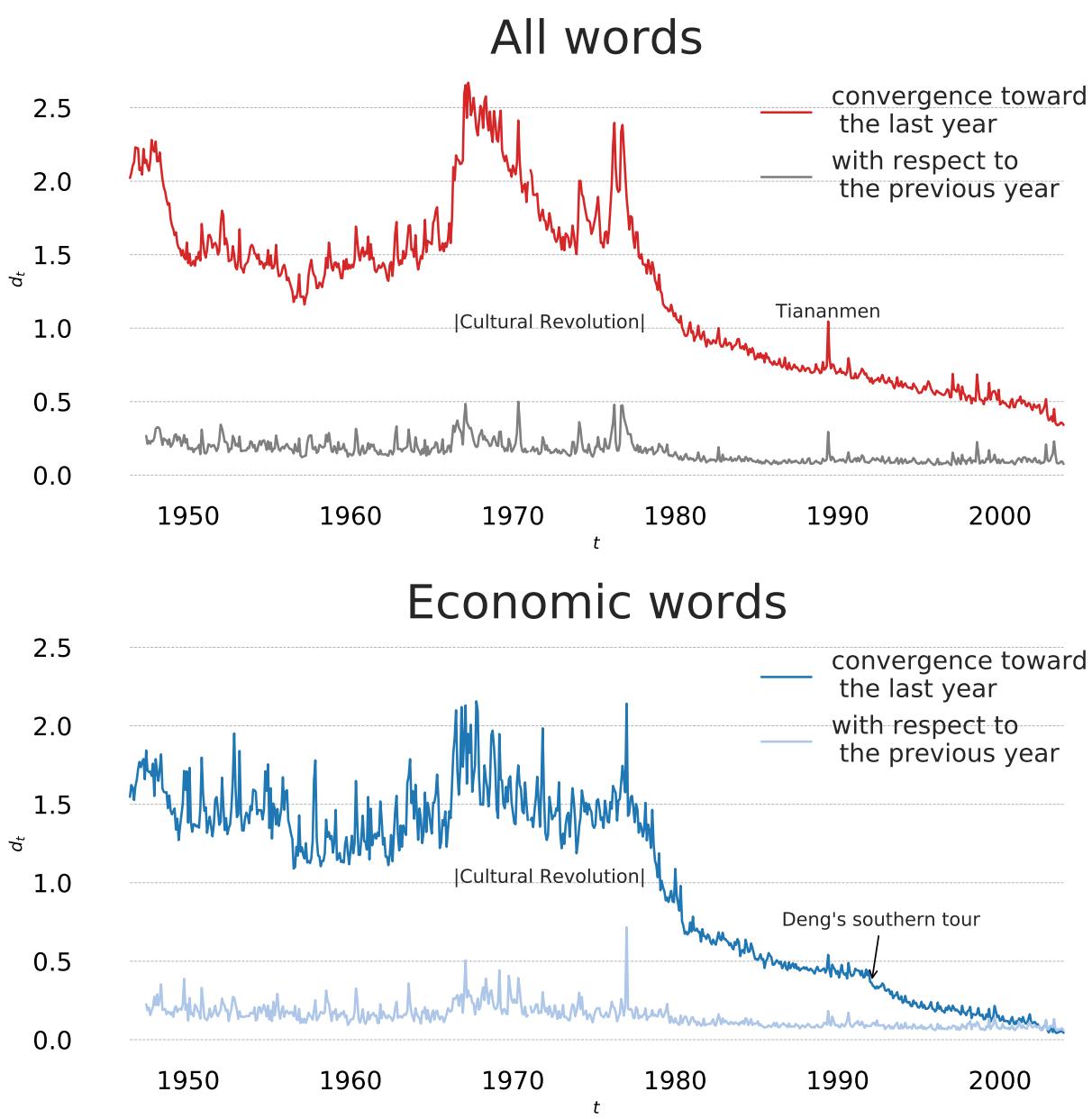


Figure 3: Monthly frequency change measured in KL divergence

In contrast to the overall stability, patterns in the economic domain were somewhat different. There was 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 4. One was applied to all economic words at all time points, and the other was 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 were similarly loaded. Arrows were 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 was needed to align all the years in a linear order. Unlike in previous years, the CCP’s economic discourse never went backwards in the last two decades covered by our analysis.

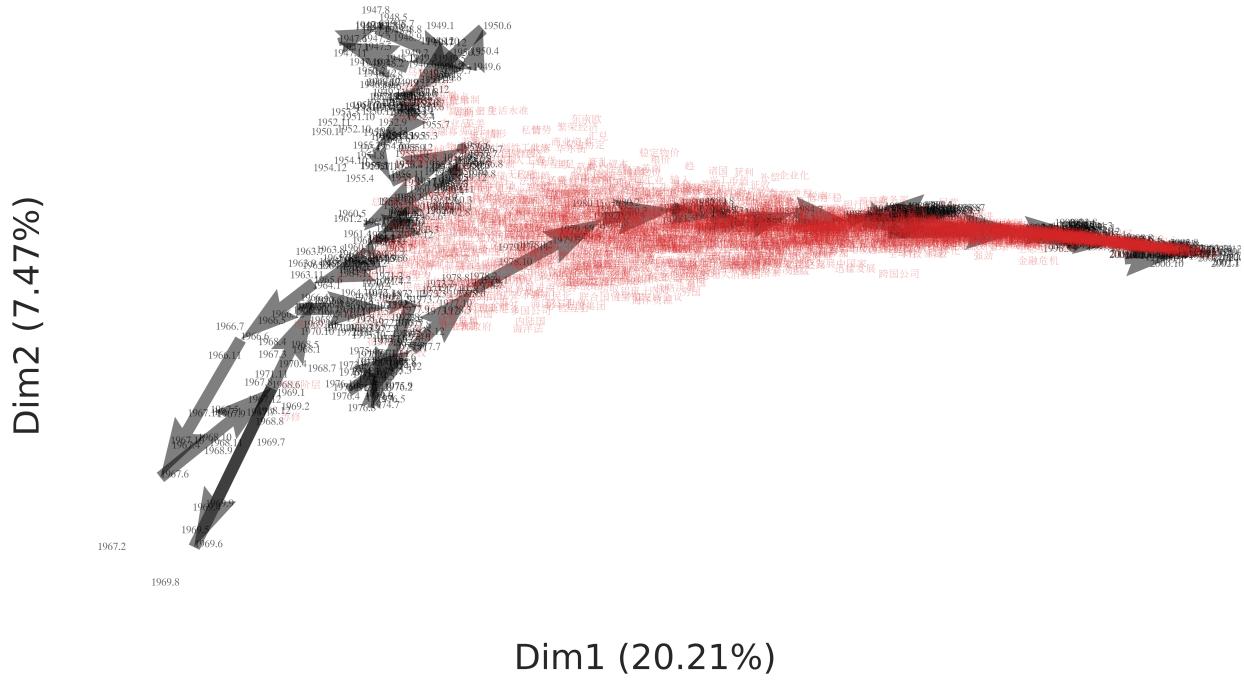
However, with the help of the second dimension, a change in discourse could again be detected. In the spring of 1992, Deng Xiaoping, who was officially retired but still the most powerful political figure, made a series of public visits to several special economic zones in South China and strongly signaled his dissatisfaction with the progress of China’s reforms. Authorities in Beijing eventually responded positively in the *People’s Daily*. The event marked the beginning of 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 trigger of change in internal patterns. More about this will be explained in the next section.

New Wine in Old Bottles

Next, we found 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 measured yearly changes in discourse in the word-embedding spaces. First, we selected the words most similar to “economy” and “politics” in the last year. Then, we extrapolated the probability of their co-occurrence in all previous years. The results, which are shown in Figure 5, were similar to what we found in the word-frequency analysis. In both the economic and political domains, there was a “recovering” process. Some scholars argue that

Economic words, 1946-2003



Economic words, 1982-2003

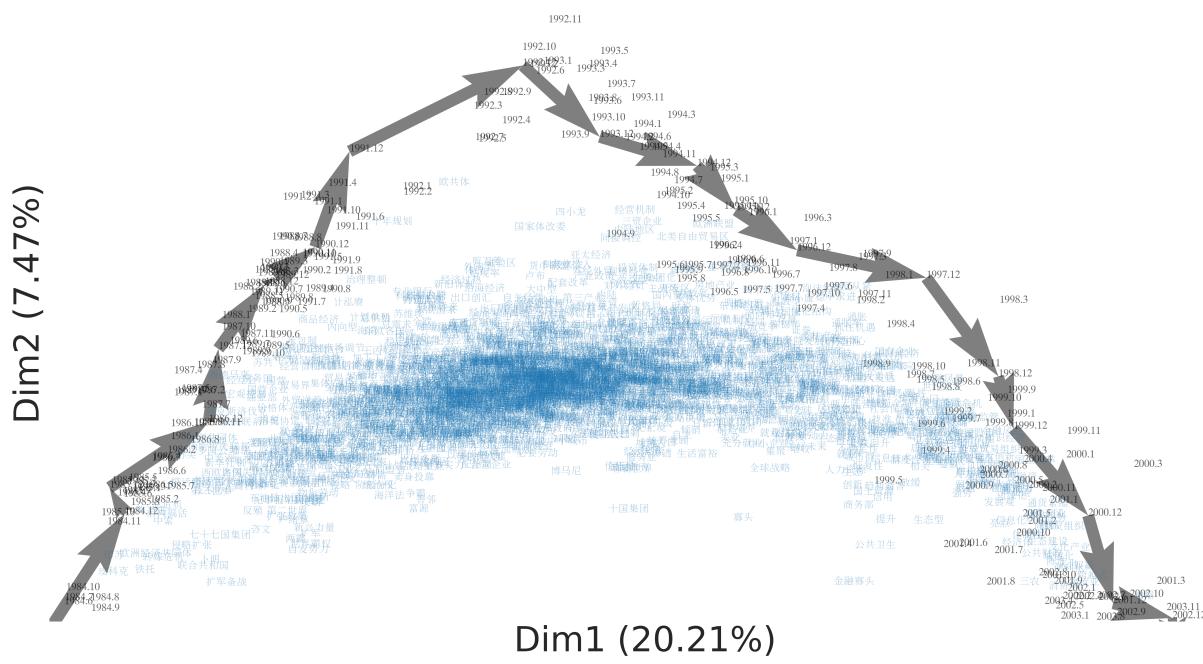


Figure 4: CA biplots in row principal components, dim 1 vs. dim 2. In the graph on the left, the first principal dimension captures an ideological shift from the leftist point during the Cultural Revolution toward the latest time point in the reform period. The second dimension captures how the Party's 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.

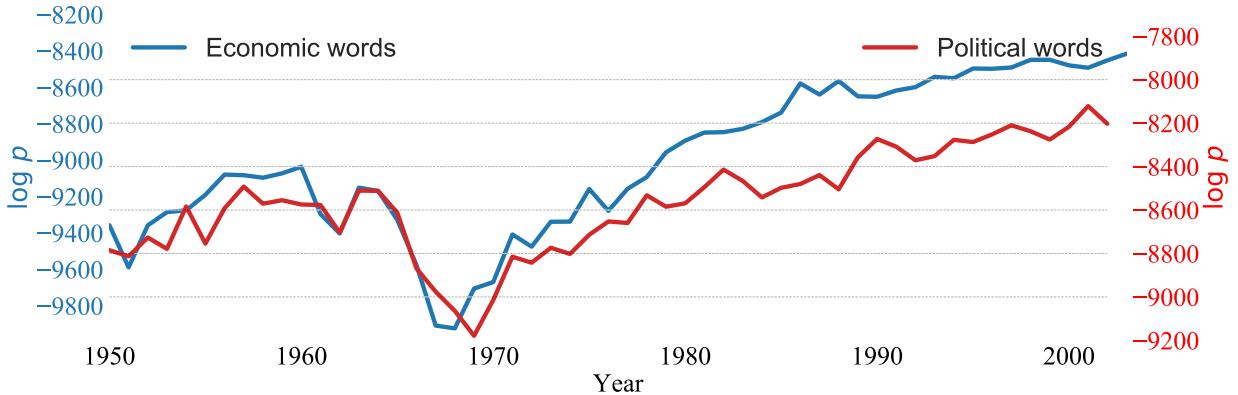


Figure 5: Log-likelihoods of co-occurrences of the most recent year’s political and economic words given each year’s corpus.

the Cultural Revolution destroyed China’s state bureaucratic machine and unintentionally paved the way for later institutional reforms (Shirk 1993; Naughton 1995). In contrast, 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).

Next, we paid close attention to what happened in 1992. We found that when a sudden change happened, newly introduced concepts still needed to be attached to existing stable repertoires to become stabilized. Figure 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 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 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 examined the causal effect of field instability on word survival by employing dynamic Poisson mixed-effects models. Estimates with three different covariate specifications

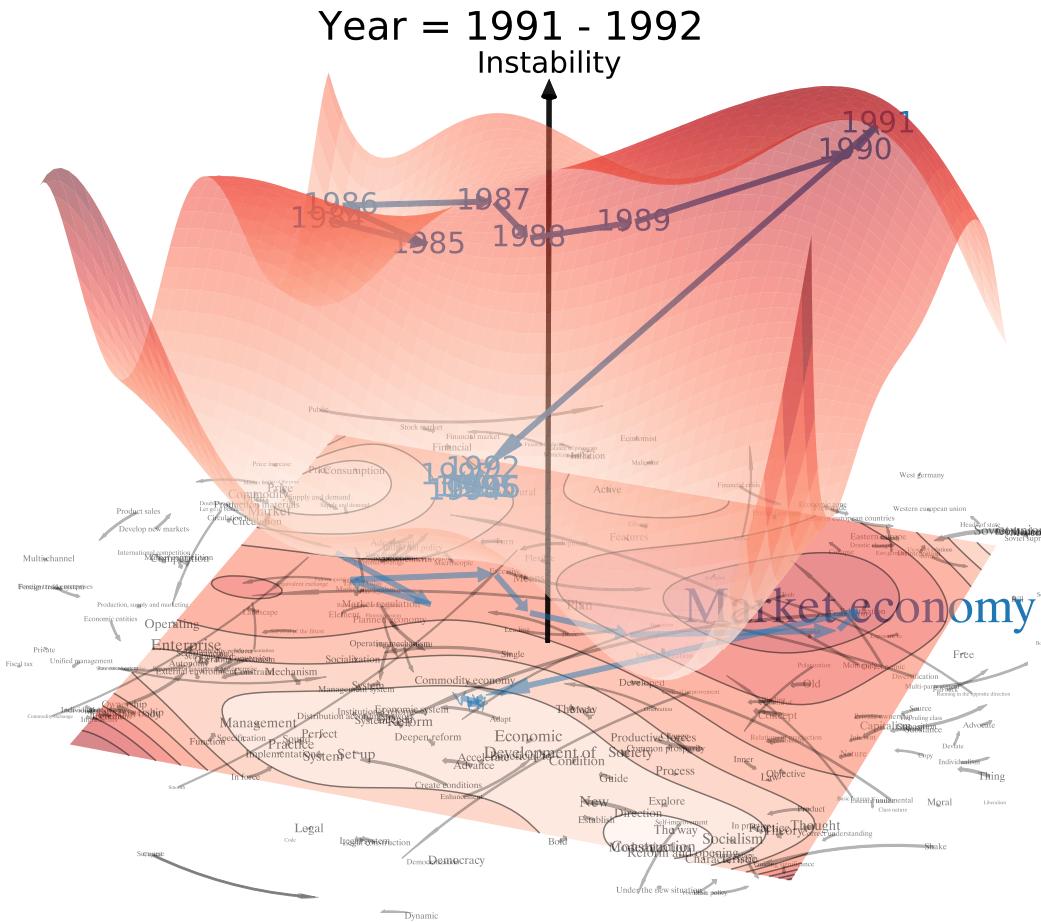


Figure 6: The xy -plane is a t-SNE (Maaten and Hinton 2008) visualization of the vector space around the word “market economy” from year 1991 to 1992. The z -axis represents the instability of the field. A polynomial surface of degree 6 was estimated based on the instability measures of the words involved. 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.

are reported in Table 1. Model (1) was a baseline model with only temporal dependence and fixed and random effects. Our focal variable, field instability, was included in Model (2). k -nearest neighbors, with $k = 50$, were used in the measurement of field instability. Model estimates with $k = 30$ and/or a different measure of field instability are reported in Supplementary Section 5. Self-instability was also controlled because it could be associated with field instability, and previous studies have found that frequency and self-instability are highly correlated (William L. 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 exerted a strong and direct negative effect on word frequency. In other words, the more unstable a word’s field was, the less likely it was 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 was controlled in Model (3), and its effect was found to be minimal; thus, we could exclude self-instability as a confounder.

To investigate further whether the field effect was truly exogenous¹, we focused on the

1. Here, we use the word “exogeneous” to mean that the treatment variable is not correlated with the error

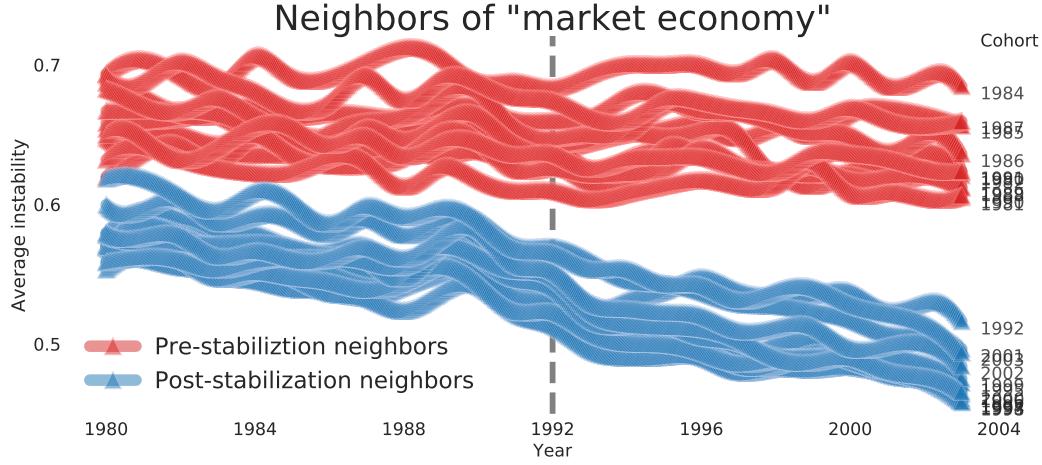


Figure 7: The yearly cohorts of the 100-nearest neighbors of “market economy” were first obtained from the word-embedding models, and their yearly average instabilities were plotted over time. It can be seen that the post-1992 neighbors were already more stable than pre-1992 neighbors before the year of stabilization. In other words, the phrase were settled in an existing stable neighborhood.

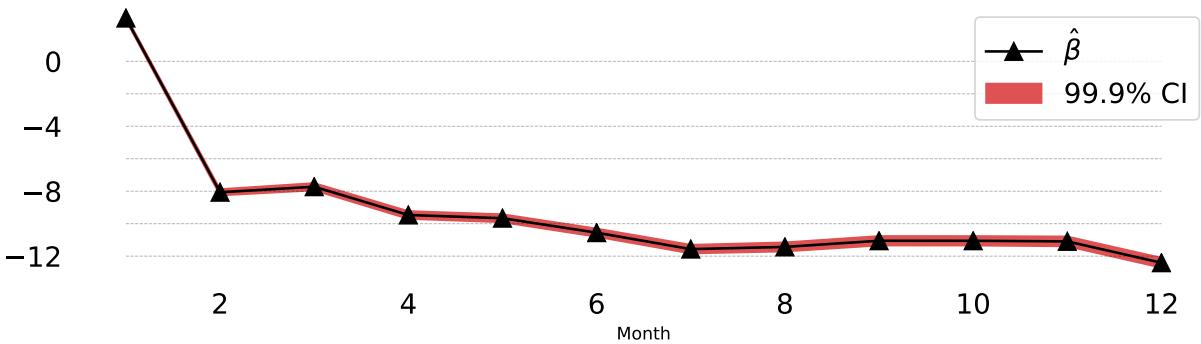


Figure 8: Effect of field instability on frequency in the first 12 months since inception

initial conditions of all the words during the first 12 months since their inception, and regressed every month’s frequency on field instability, controlling for birth cohort in a Poisson regression. Coefficient estimates are plotted in Figure 8. The results showed that the effect was positive in the 1st month and only became negative after the 2nd month. In other words, the words that were born in unstable fields had higher initial frequencies than those born in stable fields. The field stabilizing effect only came later after birth. Year-by-year results are shown in Supplementary Section 6. The pattern was consistent across all years.

Similarly, when the question of where new words were born was asked, field instability was found to be positively related to birth. Logistic regression results are reported in the last column of Table 1 to show whether a word was more likely to be a new word or old word given its field instability. In contrast to survival, birth was more likely to occur in unstable fields rather than stable fields.

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

terms. Within our theoretical framework, the field effect offers an internalist (endogenous) explanation of rhetorical change

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

Outcome	Frequency			Birth
	(1)	(2)	(3)	
Lagged log frequency	0.382*** (0.004)	0.085*** (0.005)	0.409*** (0.007)	
Self-instability		-3.340*** (0.006)		
Lagged self-instability			0.268*** (0.006)	
Field instability		-5.600*** (0.011)	-6.100*** (0.012)	6.522*** (0.092)
Initial year × Lagged log frequency	-0.067*** (0.010)	0.062*** (0.009)		
Constant	2.065*** (0.035)	9.200*** (0.040)	6.154*** (0.030)	11.149 (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.

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 2012). 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 eco-

nomic 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 classical question of “where novelty comes from.” In studies of knowledge production, it is a 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 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 analysis 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:

$$D_{\text{KL}}(P\|Q) = \sum_i P(w_i) \log \frac{P(w_i)}{Q(w_i)} \quad (1)$$

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:

$$d_t = D_{\text{KL}}(P_t\|Q_t), \quad (2)$$

where P_t is the probability distribution of words in the t th 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, \dots, w_N , the objective of the skip-gram model is to find a two-layer representation of words to maximize the average log probability,

$$\frac{1}{N} \sum_{i=1}^N \sum_{j=-k}^k \log p(w_{i+j}|w_i) \quad (3)$$

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

$$p(w_i|w_j) = \frac{\exp(\mathbf{u}_{w_i}^\top \mathbf{v}_{w_j})}{\sum_{l=1}^{\mathcal{V}} \exp(\mathbf{u}_{w_l}^\top \mathbf{v}_{w_j})}, \quad (4)$$

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:

$$\begin{aligned} \log(\lambda_{i,t}) &= \alpha + \gamma \log F_{i,t-1} + \mathbf{x}_{i,t}^\top \boldsymbol{\beta} + \mu_i + \nu_t \\ F_{i,t} &\sim \text{Poisson}(\lambda_{i,t}), \end{aligned} \tag{5}$$

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/1gkihvJ24mhB6qasU635HdprfqezFSXO/view?usp=sharing>.

References

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

- Alexander, Jeffrey C., and Philip Smith. 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.
- Li-An, Zhou. 2007. "Governing China's Local Officials: An Analysis of Promotion Tournament Model." *Economic Research Journal* 7:36–50.
- Bengio, Yoshua, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. 2003. "A neural probabilistic language model." *Journal of machine learning research* 3 (Feb): 1137–1155.
- 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.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: 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.
- Cohen, Paul A. 1988. "The post-Mao reforms in historical perspective." *The Journal of Asian Studies* 47 (3): 518–540.
- 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.
- Durkheim, Émile. 2001. *The Elementary Forms of Religious Life*. Edited by Mark Sydney Cladis. New York: Oxford University Press.

- 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.
- 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.
- 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.
- 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.
- Kane, Anne. 1991. “Cultural Analysis in Historical Sociology: The Analytic and Concrete Forms of the Autonomy of Culture.” *Sociological Theory* 9 (1): 53–69.
- Kaufman, Jason. 2004. “Endogenous explanation in the sociology of culture.” *Annual Review of Sociology* 30 (1): 335–357.
- Klingensteins, 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.
- 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.

- 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.
- Lenin, Vladimir Ilich. 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.
- Lieberson, Stanley. 2000. *Matter of Taste: How Names, Fashions, and Culture Change*. New Haven, CT: Yale University Press.
- Maaten, Laurens van der, and Geoffrey Hinton. 2008. “Visualizing Data using t-SNE.” *Journal of Machine Learning Research* 9:2579–2605.
- Martin, John Levi. 2003. “What Is Field Theory?” *American Journal of Sociology* 109 (1): 1–49.
- Marx, Karl. 1974. *The German Ideology*. London: Lawrence Wishart.
- 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.
- Naughton, Barry. 1995. *Growing out of the Plan: Chinese Economic Reform, 1978-1993*. New York: Cambridge University Press. Book.
- Nee, Victor. 1989. “A theory of market transition: From redistribution to markets in state socialism.” *American Sociological Review*, 663–681.
- Oi, Jean C. 1999. *Rural China Takes Off: Institutional Foundations of Economic Reform*. Berkeley: University of California Press.
- Padgett, John F, and Walter W. Powell. 2012. “From chemical to social networks.” In *The Emergence of Organizations and Markets*, edited by John Frederick Padgett and Walter W. Powell, 92–114. Princeton: Princeton University Press.
- Palla, Gergely, Albert-László Barabási, and Tamás Vicsek. 2007. “Quantifying social group evolution.” *Nature* 446 (7136): 664.

- Perry, Elizabeth J. 2012. *Anyuan: Mining China's Revolutionary Tradition*. Berkeley: University of California Press.
- Polanyi, Karl. 2001. *The Great Transformation: The Political and Economic Origins of Our Time*. Boston: Beacon Press.
- 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 reserving market incentives." *Journal of Economic perspectives* 11 (4): 83–92.
- Rule, Alix, Jean-Philippe Cointet, and Peter S Bearman. 2015. "Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014." *Proceedings of the National Academy of Sciences* 112 (35): 10837–10844.
- Saussure, Ferdinand de. 2011. *Course in General Linguistics*. New York: Columbia University Press.
- Shirk, Susan L. 1993. *The Political Logic of Economic Reform in China*. Berkeley: University of California Press. Book.
- 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.
- 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*.
- 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.
- Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* 342 (6157): 468–472.

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

Walder, Andrew G. 1995. “Local governments as industrial firms: An organizational analysis of China’s transitional economy.” *American Journal of Sociology* 101 (2): 263–301.

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

Weber, Max. 1948. “Religious rejections of the world and their directions.” In *From Max Weber*, edited by H. H. Gerth and C. Wright Mills, 323–359. London: Routledge & Kegan Paul Ltd.

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

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, Mass.: Harvard University Press.