

NBER WORKING PAPER SERIES

IDEAS MOBILIZE PEOPLE:
THE DIFFUSION OF COMMUNIST IDEOLOGY IN CHINA

Ying Bai
Ruixue Jia
Runnan Wang

Working Paper 30947
<http://www.nber.org/papers/w30947>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2023

No relevant financial relationships. We thank Sam Bazzi, Harris Doshay, Matthew Gentzkow, Marvin Suesse, Yang Yao, Enying Zheng and seminar/conference participants at National School of Development (PKU), CUHK-Shenzhen, Academia Sinica, Trinity College Dublin, PKU-Tsinghua conference on Chinese economy, and UCSD for their comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Ying Bai, Ruixue Jia, and Runnan Wang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Ideas Mobilize People: The Diffusion of Communist Ideology in China
Ying Bai, Ruixue Jia, and Runnan Wang
NBER Working Paper No. 30947
February 2023
JEL No. D70,D71,D83,N45,O12,P30

ABSTRACT

Can ideas mobilize people into collective action? We provide a positive answer to this question by studying how exposure to the Communist ideology shaped an individual's choice to join the Chinese Communist Party (CCP) during the party's formative stage. The individuals we focus on are cadets at the Whampoa Military Academy, who subsequently fought in 20th-century China's most important wars. Our identification strategy exploits the locality-time-content variation in the circulation of the *New Youth* magazine—the major platform to promote Communism after the Treaty of Versailles in 1919—as well as the variation in an individual's location over time. By comparing the Whampoa cadets living in a locality with post-1919 *New Youth* available against those who had lived in the same locality but missed this channel, we demonstrate that the former were significantly more likely to join the CCP. In future political struggles, those whose party choice was more influenced by this ideology channel were less likely to quit the CCP and more likely to sacrifice their lives. Additionally, we document that family background cannot predict the party choice of these political pioneers but social networks can complement ideology exposure.

Ying Bai
Department of Economics,
The Chinese University of Hong Kong
Shatin, Hong Kong
rickybai@gmail.com

Runnan Wang
Department of Economics
The Chinese University of Hong Kong
Shatin, Hong Kong
rwan@link.cuhk.edu.hk

Ruixue Jia
School of Global Policy and Strategy
University of California at San Diego
9500 Gilman Drive #0519
La Jolla, CA 92093
and CEPR
and also NBER
ruixue.jia@gmail.com

1 Introduction

How do people get mobilized into collective action that sometimes changes the trajectory of history? Among many determinants, the importance of ideas has long been noted and debated. Do new ideas cause political and economic actions? Or is it that such ideas are more-so used to cover one's own material incentives?¹ Perhaps because the impact of ideas is often elusive and difficult to prove, there has been little empirical investigation into the relevance of ideas in social and political movements. In this study, we use the spread of Communist ideology in China to demonstrate how ideas can affect individuals' political actions, which ultimately would shape the fate of modern China.

The spread of Communist ideology was one of the most significant political events of the 20th century: For a large part of the century, approximately one-third of humanity lived in Communist countries. Yet, little is known about how such an ideology got diffused across such vastly different contexts ranging as Germany and Italy in the west and China and Indonesia in the east. China, during the 1910s–1920s, provides a useful context for Communism's spread via the printing press. As well recognized by abundant historical narratives, the most prominent press in spreading communist ideology is the *New Youth* magazine (1915–1922), a leading literary magazine founded in 1915 that initiated the New Culture Movement (1915–1923) but turned to promoting Communism after the Treaty of Versailles in 1919.²

Identifying the impact of ideology diffusion, however, is challenging. First, it is difficult to differentiate ideological exposure from other omitted variables. For instance, the circulation of the *New Youth* magazine across localities would, in most cases, be associated with locality size or urbanization rates. Second, it is similarly hard to determine the direct population that was influenced by the ideology. In a setting like the one being discussed in this paper, where over 80% of the population were illiterate and printing press was the primary form of media, it is conceivable that

¹More broadly, the debate between idealism and materialism as an explanation for historical change has been going on since the early days of civilization. Among modern classical works on ideas, [Hirschman \(1977\)](#) discusses how the ideas of capitalism initially emerged. [Berlin \(1979\)](#) emphasizes the importance of dissenters in the history of ideas.

²Even though China was on the winning side of WWI, the Treaty of Versailles allowed Japan to retain territories in Shandong that had been surrendered to Germany. Such actions of Western powers undermined the moral prestige of the West in China and directed the leadership of the *New Youth* towards Russia for enlightenment. See more on the background in Section 2.

only a small share of the population could be influenced. Although these individuals might go on to become pioneers in the coming Communist revolution, it is challenging to pinpoint who exactly they were.

We are able to overcome these challenges by exploiting a specific context and important dataset: the choice of individuals at the Whampoa Military Academy to become a Chinese Communist Party (CCP) member during the CCP's formative stage in the 1920s. The Whampoa Military Academy was established in 1924 and produced commanders and legendary graduates who would subsequently go on to fight China's most important wars of the 20th century, including two civil wars between the CCP and the Kuomintang (KMT) as well as the anti-Japanese war. The first Whampoa cohort included 689 individuals who provided detailed information on their family background, life history, and even the references they used when applying to the Whampoa. Because these were important individuals, we are also able to gather information on their future political outcomes. Their rich biographical information provides a laboratory to study political choice at the individual level in an important historical setting, a task that has proved to many who have come before.

Specifically, we examine how the exposure to Communist ideology via the *New Youth* affected an individual's decision to become a CCP member. Our identification strategy leverages two variations. One is a time-varying geographical variation in the circulation of the *New Youth* and its content change before and after 1919. The other is a time-varying location of individuals (due to education or employment). We compare cadets who were living in a locality with the post-1919 *New Youth* available against those who had been to the same locality but missed the post-1919 *New Youth* due to the timing of their life events. In other words, we can address the challenge that ideology exposure is correlated with other locality characteristics by controlling for locality fixed effects. Our analyses use two definitions of localities—prefectures and counties³—for comparison and robustness. We show that, conditional on locality fixed effects, individuals with or without access to post-1919 *New Youth* are comparable in terms of a rich set of personal characteristics, including family socioeconomic status, education, the political identities of one's references, age, numbers of siblings and children, marriage, and even behavioral traits such as smoking and drinking. Therefore, given the same locality, exposure to the Communist ideology is close to being random. This pattern suggests that an individual's location is primarily driven by economic concerns (especially education and jobs) rather than the presence of the *New Youth*.

We find that exposure to the *New Youth* at the prefecture during 1919–1922 increases an

³A prefecture is a collection of counties. Roughly speaking, one prefecture consists of approximately 10 counties.

individual's probability of joining the CCP by 11.7 percentage points when, compared with those who lived in the same locality but missed this exposure. The magnitude becomes 17.2 percentage points when we zoom into county-level exposure. These estimates imply a persuasion rate of at least 18.9%. This relatively large persuasion rate is partly because our studied sample is a group of educated young radicals (whose average age was 18 in 1919) on which ideology exposure would likely be particularly important.⁴ In contrast, the other personal characteristics, except for the political identity of one's references, cannot predict the decision to join the CCP and the KMT and do not offer strong heterogeneous impact for ideology exposure. The insignificant impact of economic status is partly because the CCP was at its inception stage and faced a great deal of risk.

Two main concerns of our main finding are whether the *New Youth* entered certain localities due to potential demand and whether individuals relocated to those localities in response to the availability of post-1919 *New Youth*. We should note that our estimate is similar even if controlled for the interactions between prefecture fixed effects and the post-1919 dummy, which allows us to then control for possible changes in the demand for Communism in different prefectures before and after 1919. Moreover, we take a closer look at possible endogeneity concerns regarding the circulation locations of the *New Youth* and the relocation of individuals. First, we find that exposure to the *New Youth* during 1915–1918 had no similar influence on the CCP choice, suggesting that our finding is not driven by the endogenous entry of the *New Youth*. Second, our estimate is similar when restricting exposure to the localities that *always* had access to the *New Youth*, which helps eliminate the possible entry concerns of *New Youth* after 1919. Third, to more effectively capture the location preferences for an individual, we construct placebo locations of an individual by assuming that he were in his previous or following location based on his life event records. We use these placebo locations to calculate placebo exposures to post-1919 *New Youth* and find that these placebo exposures cannot explain our finding either. Finally, we separate our comparison group into two subgroups: (i) those who had never been to localities with post-1919 *New Youth* and (ii) those who had been to localities with post-1919 *New Youth* available but missed the magazine. We find an estimate similar to our baseline when comparing with group (ii) whose relocation is not driven by the presence of post-1919 *New Youth*. We discuss concerns revolving around possible measurement error at both the locality and individual levels and show that our findings are unlikely to be driven by measurement error at either level.

We focus on post-1919 *New Youth* exposure because it is widely recognized as the main

⁴We also explore very local exposure based on school information and obtain an upper bound of 29.2%. DellaVigna and Gentzkow (2010) summarize persuasion rates across different settings, ranging from 0.7% to 30%.

channel for Communist ideology diffusion (see qualitative evidence in Section 2). It is certainly possible, however, for our studied individuals to read other magazines or be exposed to activist groups in this period, some of which might also help spread Communist ideology. Because of this, we conduct three sets of analyses to help interpret post-1919 *New Youth* exposure. First, based on archival information on 19 pro-Communism periodicals that existed during 1919–1922, we are able to characterize a network by examining whether these magazines shared editors or editorial offices. This network reveals that the *New Youth* was the most central magazine and its editors often would be instrumental in helping co-found other magazines. Second, comparing the exposure to post-1919 *New Youth* to that of two other pro-Communism magazines, we find that few would have been exposed to the other magazines without being exposed to post-1919 *New Youth*. Estimates are similar regardless of whether we further consider these additional periodicals. Third, we use archival information on activist groups in this era and find our result still holds in localities without these groups. There does exist some complementarity between the ideology exposure we focus on and the presence of these activist groups. Based on these analyses, we interpret post-1919 *New Youth* exposure as a reasonable proxy for the diffusion of Communism ideology in a locality.

In the last part of our analyses, we examine whether these individuals that have been mobilized by ideas behave differently in critical historical moments. This is an important yet challenging issue to study. To explore this issue we investigate the future political outcomes of our studied individuals through various political struggles, during which many quit their parties and sacrificed their lives. We find suggestive evidence that those who were more influenced by the Communist ideology exposure channel were less likely to quit the CCP and more likely to sacrifice their lives. Moreover, this difference is primarily driven by the period from 1927 to 1936, during the first Civil War between the CCP and the KMT, when the CCP was the weaker party and was under severe attacks by the KMT. These findings suggest that initial ideology exposure can work alongside other factors of the revolution process to lead one toward significantly different political choices.

Our study contributes to evaluating the importance of ideas in shaping political and economic processes, which has long been recognized but is difficult to quantify.⁵ In a series of perceptive works, Joel Mokyr points out the importance of a set of ideas collectively known as “the Enlightenment” in creating the Industrial Revolution (e.g., Mokyr 2005). A burgeoning literature in economics has explored the impacts of ideological exposure on voting (e.g., DellaVigna and Kaplan 2007),

⁵A long-standing example of ideas is religion. Existing literature has focused more on how material incentives affect religion diffusion (e.g., Cantoni 2012, Michalopoulos, Naghavi and Prarolo 2018, Becker and Pascali 2019) than the impacts of ideas separate from material incentives.

public opinion (Cantoni et al. 2017) and policy-making (Ash, Chen and Naidu 2022). Evidence on the relationship between the significance of ideas and large-scale political and social movements, however, is scarce. To the best of our knowledge, our study is the first to trace the diffusion of Communism ideas and demonstrate its impacts on an individual's political choices.⁶

We are also able to compare the role of ideology exposure with that of family background and proxies for social networks. We find that one's political choice is positively associated with his personal references' political identities, consistent with a recent literature that highlights the importance of social networks (e.g., Satyanath, Voigtl nder and Voth 2017, Enikolopov, Makarin and Petrova 2020).⁷ This paper also documents some complementarity between social networks (proxied by the identity of references and the presence of activist groups) and ideology exposure. In contrast, family background is not predictive of CCP membership. Our research can be extrapolated to other scenarios where political or social movements face a great deal of risk, and in turn, better understand how the ideas and social influences around these movements are likely to be particularly relevant.⁸

In addition, our studied individuals would become leaders in the Communist revolution. While a growing literature has documented the importance of leadership for economic and political development (e.g., Jones and Olken 2005, Dippel and Heblich 2021, Cag , Dagherret, Grosjean and Jha 2022, Bai, Jia and Yang 2022), we do not know much about where leaders come from. Our findings suggest that ideology can shape leaders and their behavior in critical moments of history.

Finally, scholars studying the origins of the CCP have recognized the formative role of the *New Youth* (e.g., Dirlik 1989, Meisner 1999, Smith 2000, Ishikawa 2013). However, due to the empirical challenges mentioned above, the influence of the *New Youth* remains a historical conjecture. Our study provides an approach to trace the spread of the *New Youth* and evaluate its impact on a group of individuals who would jointly determine China's political path. Although we focus on China, the role of the printing press in spreading Communist ideology is likely to be relevant in other contexts.

⁶There exist many volumes of historical studies on Communism, although focusing on a few nations as case studies (Smith 2014). Magness and Makovi (2022) document the Russian Revolution's central role in mainstreaming the ideas of Marx. In our setting, the Russian Revolution and the Treaty of Versailles facilitated the adoption of Communism in China.

⁷In our context, this pattern is nicely summarized as "from friend to comrade" in historical narratives (Van de Ven 1992). It should be noted that there is an extensive literature in political sociology on the determinants of participation in social movements which has discussed social networks (see Walder (2009) for a review).

⁸Those who joined the CCP after CCP grew stronger could be very different from the political pioneers we focus on. For instance, some scholars argue that nationalism against Japan might have contributed to the growth of the CCP after 1937 (Johnson 1962, Chen and Kung 2020).

From this lens, our study joins a growing literature on the impact of media in shaping political behaviors. The existing literature has documented the influence of various forms of media in a variety of contexts (see DellaVigna and Gentzkow (2010), Prat and Strömberg (2013) and Zhuravskaya, Petrova and Enikolopov (2020) for reviews), most of which focus on macro-regional patterns.⁹ Our setting provides individual-level evidence while also highlighting the role of magazines, a medium whose influence has not been paid much attention. Moreover, beyond demonstrating the causal impact of ideology exposure on political actions, we find suggestive evidence that such an impact may also partly determine the life and death of political pioneers in the long run.

2 Background and Data

2.1 The *New Youth* (1915–1922) and the Founding of the CCP in 1921

The period of 1910s–1920s, known as the “Chinese enlightenment” (Schwarcz 1986), is exemplified by the New Culture Movement (1915–1923), the later stage of which is also known as the May Fourth Movement (1919–1923).¹⁰ Arising out of a disillusionment with traditional Chinese culture after the failure of the Beiyang government (1912–1928) to address social and economic problems, many intellectuals joined the movement to promote a new culture based upon western ideals, particularly democracy and science. The movement was launched by the writers of the *New Youth*, a monthly magazine founded in 1915 by Chen Duxiu. Different from those who thought that democracy and constitutional government could be easily established in China, Chen argued that there must be a change in *thought* to support the constitutional government (Chen 1916). He believed that there couldn’t be a new political system without a new culture.

Following the emphasis on culture, the *New Youth* magazine initially focused on culture and literature, offering such debates as promoting written vernacular Chinese over classical Chinese. It became popular very quickly with its monthly circulation rising from 1,000 for the first issue to over 16,000 after 1917 (Chow 1960). Over time, frustrated with the inadequacy of cultural change and influenced by the Russian Revolution, the magazine, led by Chen Duxiu and Li Dazhao, started

⁹As an important example in this literature, Adena et al. (2015) use the variation in the locality-slant variation in radio exposure to demonstrate how media exposure contributed to the rise of the Nazis across localities in prewar Germany.

¹⁰The May Fourth Movement is named after the May Fourth incident on May 4, 1919. On that day, students gathered in front of Tiananmen to protest the government’s weak response to the Treaty of Versailles. These demonstrations sparked nationwide protests and contributed to the demand for political action.

considering more political action. The 1919 Treaty of Versailles marked a turning point, where Chinese intellectuals, disillusioned at the compromising of Chinese interests by Western powers, found the Russian way as a new role model.¹¹ In May 1919, the *New Youth* published a special issue on Communism. After that, the magazine became increasingly aligned with the Communist ideology and served as the fountainhead of Chinese Communism (Ishikawa 2013).

To track this shift in ideology, we plot the share of *New Youth* articles titled with key phrases in Figure 1. Panel (a) focuses on the mention of different nations. Panel (b) focuses rather on selected terms and topics. When *New Youth* started in 1915, France and Britain were the most frequently referenced nations, reflecting a promotion of western ideals. In contrast, after 1919, Russia became the most discussed, accounting for 10–15% of all articles published in a given year. Alternately, the culture/literature topic dominated initially but political discussions related to Communism and revolution dominated after 1919–20. In Appendix A.1, we further plot the most frequently used words in titles. As shown, after 1919, words related to Communism (e.g., labor, Russia, socialism, Marx, and revolution) stood distinctly out. As reflected by these words, Communism can be considered as a set of ideas related to a proletarian revolution that involves getting organized for “unified action” (Dirlik 1989).

The leaders of the *New Youth* were also among those who founded the CCP. In July 1921, the first National Congress of the CCP was held in Shanghai. It was here that Chen Duxiu was elected as the Chief Party Secretary, and following the event, the *New Youth* would then become a theoretical journal for Communism. The *New Youth* published its last issue in July 1922.¹² The *New Youth* published 54 issues in total during 1915–1922, which has served as an important source for studying the political intellectual history of modern China.¹³

Abundant anecdotal evidence exists on the importance of the *New Youth* in shaping the ideology of the young in this era, including memoirs from future leaders of the CCP. For instance, Mao Zedong was one of the avid readers and made his first contribution to the magazine in 1917. In his 1936 interview with Edgar Snow (Snow 1937), Mao reminisced:

“I began to read this magazine [*New Youth*] while I was a student in the normal school

¹¹In sharp contrast with the Treaty of Versailles, the Karakhan Manifesto, issued in 1919 on behalf of the Soviet government, nullified unequal treaties forced upon China by Czarist Russia. Not surprisingly, Russia “ascended the ideological throne forfeited by the West, and became the symbol of justice and freedom” (Huang 1964).

¹²Although the original *New Youth* was stopped in 1922, the CCP continued to use the title for a new quarterly journal during 1923–1926. Because of the timing of our study, we do not focus on this latter period.

¹³For instance, Jin and Liu (2008) analyze the word frequencies in the *New Youth* and confirm the magazine’s transition to being a political journal after 1919.

and admired the articles of Hu Shi and Chen Duxiu very much. They became for a while my models, replacing Liang Qichao and Kang Youwei.”¹⁴

Similar recognitions of the formative role of the *New Youth* can be found in the self-reflections of other Communist pioneers including in Zhou En-lai’s discussion with Edgar Snow (Snow 1937), Yun Daiying’s diaries (Yun 1981), Deng Xiaoping’s 1926 autobiography (Yang 1994), and Shi Cuntong’s memoir (Yeh 1996). Notably, these individuals were located in various regions, indicating that the influence of the *New Youth* went beyond metropolitan cities such as Beijing and Shanghai.

Given the importance of the *New Youth*, few scholars studying the history of the CCP have missed mentioning it. As put by Meisner (1999), “it is difficult to overestimate the importance of the intellectuals who coalesced around the *New Youth*, for their writings molded the beliefs and attitudes of a whole generation of young students who were to achieve political prominence after the May Fourth incident of 1919 and who were to become the leaders of the modern Chinese revolution.”

The post-1919 *New Youth* was the major but not the only channel for Communist ideology diffusion. Following in the footsteps of the *New Youth*, quite a few pro-Communism periodicals emerged, some of which were co-founded by the editors of *New Youth*. Moreover, there existed various activist groups that also helped spread Communism. We study how these alternative channels affect interpreting the role of the *New Youth*.

2.2 The Whampoa Military Academy (1924–1927)

The Whampoa Military Academy was established in 1924 in Whampoa, Guangzhou, as proposed and supported by the Soviet Union. In 1924, under the demand of the Soviet Union, the KMT formed a political alliance with the CCP to fight against the incumbent Beiyang government and other regional warlords. In this era, known as the First United Front, the KMT was the dominant revolutionary party whereas the CCP was at its more nascent stage. While initially everyone at Whampoa was required to be a KMT member, one could choose to join the CCP (if they hadn’t before 1924). This KMT-CCP collaboration, however, broke apart in 1927, marking the beginning of a civil war that would last until 1937. After the breakup, one was only able to have one party identity. During 1937–45, the KMT-CCP formed a second political alliance (i.e., the Second United Front) to fight against Japan. After winning the anti-Japanese war, the KMT and the CCP fought in

¹⁴Liang Qichao and Kang Youwei were influential intellectuals who proposed reforms (instead of revolution) of the imperial system.

another civil war until the CCP won the mainland and the KMT withdrew to Taiwan in 1949–1950.

Despite its relatively short presence, the Whampoa Military Academy played a profound role in Chinese modern history.¹⁵ Whampoa concentrated the revolutionary military talents into training regiments that followed the Soviet Red Army model. The exigencies of the revolution limited the time available for training, lasting only half a year for most cohorts. The emphasis was on the practical knowledge and skills required in the battlefield (Graff and Higham 2012). Whampoa cadets would become commanders for both the KMT and the CCP and fight both the civil wars and the anti-Japanese war.¹⁶

Moreover, the Whampoa graduates have influenced both parties' policies and governance. For Chiang Kai-shek, the leader of the KMT, the “Whampoa clique” was a pivotal force for his military organization and governance (Taylor 2009). For the CCP, when it built its first Red Army in 1927, most of its commanders were from Whampoa. In the following two decades, the CCP trained its army in the Whampoa way, which would finally serve in bringing the CCP victory.¹⁷ In 1955, the CCP ranked the Ten Founding Marshals of the People's Republic of China, and five out of ten were affiliated with Whampoa.

2.3 The Timeline

We provide a timeline of the key events during the 1910s–1949 in Figure 2. Our study focuses on how the exposure to Communist ideology via the *New Youth* during 1919–22 affected the determination to become a CCP member in the CCP's formative stage. As Figure 2 shows, the number of CCP members was initially limited, but grew dramatically in this era: There were only 432 registered Communist members at the 3rd National Congress of the CCP (June 1923), but this number rose to over 57,000 at the 5th National Congress (April 1927). In addition, we explore future outcomes (1920s–1949) of the individuals in our study to shed light on the extended impacts of ideology exposure.

By plotting the founding years of Communist parties in the various countries featured in Figure 2, along with the year of the translation of the *Communist Manifesto*, we are able to see that many

¹⁵After the KMT-CCP breakup, the KMT established the Republic of China Military Academy in Nanjing.

¹⁶For instance, legendary commanders on the KMT side include Chen Cheng, Du Yuming, Xue Yue, Hu Zongnan, Hu Lien, and Guan Linzheng, whereas the CCP side includes Lin Biao, Xu Xiangqian, Liu Zhidan, and Chen Geng.

¹⁷In a 1937 interview with James Bertram, Mao Zedong emphasized that the Red Army had inherited and advanced the traditions and practices of the Whampoa.

new Communist parties emerged in the wake of the Russian Revolution. Thus, the origin of the CCP is one of many examples that were facilitated by both international politics and domestic tensions. It is worth noting that the Chinese version of the *Communist Manifesto* was published in 1920, which is much later than many other countries and highlights that the ideas of Communism were new to China in the near-perfect synchronicity with our timeline of 1919–1920.

2.4 Data

We present the summary statistics in Appendix A.2 and the detailed data construction process in Appendix A.3–A.4. Below, we describe our key variables.

The First Cohort of Whampoa. The first cohort at the Whampoa Military Academy admitted 706 individuals, among which 17 were missing information.¹⁸ We focus on all 689 individuals with information. Among them, 530 were directly recruited by the Whampoa and were required to provide detailed information on their family background and life history at the time of enrollment in 1924. At Whampoa, they were assigned to be Teams 1–4. Appendix A.3 presents an example of the life history records, based on the *Detailed Survey for the First to Fourth Teams of the First Cohort of the Military Academy* (Central Executive Committee of the Kuomintang 1924). The other 159 cadets were recruited from a regional military school in Guangzhou. We obtain information on the background of all 689 individuals from the work of Chen Yuhuan, a historian and Whampoa expert (Chen 2017). In our analyses, we use all 689 individuals and present results restricted to Teams 1–4. We provide an example of our coding system to track life events in Appendix A.4.

We should note that not everyone tracked their life events year by year. With these datasets, there are two types of missing information. First, some omitted the ending year of their life events. Second, some reported certain life events, but did not specify the year. For the former, we assume that a life event lasted until the starting year of the next event; for the latter, we infer the years based on their birth year, record-end year, and the known elements about the school system during 1912–1922 (Fan and Popkewitz 2020).¹⁹ To check the validity of our inference, we present results by pretending that we do not know the specific years and applying the same inference method to all our sample. We also conduct checks regarding the precision of life events by examining how the

¹⁸These individuals did not appear at the Whampoa or potentially left before graduation.

¹⁹This system is known as the *Renzi Guichou* school system, referring to the fact that it was implemented during 1912 (the year of *Renzi* in the Chinese calendar) and 1913 (the year of *Guichou*). This system lasted until 1922.

number of records affects our findings.

CCP Membership and Future Political Outcomes. The main outcome of our analyses is CCP membership. The individuals we study would go on to play important roles in the decades following their Whampoa study, which means that there exists systematic information on their political party identification and career trajectories, even including whether one chose to quit their party or if they died in the political struggles between the KMT and the CCP. We code these variables collected from [Chen \(2017\)](#) and validate the data via Wikipedia and Baidu. Although it is possible for an individual to have hidden his CCP membership in the 1920s, this is not a concern for our study as our information on CCP membership is based on information disclosed later.

In the 1920s, the CCP was much weaker than the KMT. On average, 17.1% of our sample joined the CCP, and most (77%, 91 out of 118) of these CCP members joined before 1925.²⁰ 24.1% of our studied cadets died in future conflicts and 16% quit their parties (this includes surrendering to the other party in conflict). These numbers highlight the risk and uncertainty of the revolution.

Personal Characteristics. We consider 13 variables of a Whampoa cadet's personal characteristics categorized into three groups, which together provides a clear portrait: These were a group of men in their early 20s in 1924 who had received middle-school or secondary-school education and came from families of the middle and lower strata.

(i) Education and socioeconomic background. As reported in [Appendix A.2](#), 96% of the cadets had completed middle school and 61% attended secondary vocational schools. 11% had some college education or attended some college classes but usually did not hold college degrees. 54% came from families with some education or business tradition; 39%, 58%, and 3% ranked their family status as “poor”, “middle-income”, and “rich”. We code this status as a categorical variable (1–3), which has a mean value of 1.63.

Alongside education and socioeconomic background, we know the political identity of the personal references that cadets used in their application to the Whampoa and can measure how many CCP references one had. This variable proxies whether one had direct or indirect connections with existing CCP members and can be considered as a crude measure of social networks. This variable ranges from 0 to 5 and has a mean value of 0.18, implying that few could directly access CCP references by 1924.

²⁰By the 4th National Congress of the CCP (January 1925), there were 994 registered Communist members. Our sample accounts for 9.2% of all registered CCP members.

(ii) Demographic information. The average age of the studied individuals was 23 in 1924. The average number of siblings they had is 3.4. 54% of them were married and 14% had children, reflecting marriage at an early age in this era.

(iii) Religion and behavior traits. 11% reported believing in some religion (mostly Confucianism and Buddhism). 5% reported a habit of smoking or drinking.

Exposures to the *New Youth* and Communist Ideology. Each year, the *New Youth* provides information on its circulation location in at least one issue, which we use to construct prefecture-year and county-year datasets on the availability of the *New Youth*. The variation in circulation is driven by both the expansion effort of the magazine and the censorship of the Beiyang government. We present an example of this circulation data and a map of circulation locations over time in Appendix A.4 (Step 1). By 1920, the *New Youth* spread to 53 prefectures, covering 37.9% of China's population. See the maps in Appendix A.4 for the circulation locations. Notably, the *New Youth* was always available for a group of prefectures whereas there existed entry and exit in others. We leverage such a variation to examine whether the potential endogeneity concern of the *New Youth* circulation matters for our finding.

Our measure of the *New Youth* exposure comes from analyzing the interaction between the locality-year circulation variation and the variation in an individual's life history. In Appendix A.4 (Step 2), we explain our life history coding process and present maps of cadets' locations over time. By 1924, the individuals we study had been to 160 prefectures, covering 80.0% of China's population.

Considering that A and B lived in the same locality, Person A could have been exposed because he happened to be there during the years the *New Youth* was available, whereas Person B could have not been present during the years it was available. We further illustrate the interaction between cadets' locations and *New Youth* locations in our data in Appendix A.4 (Step 3). Further, our measure of Communist ideology exposure considers the ideology shift of the *New Youth* before and after 1919. If an individual was living in a prefecture (or county) when post-1919 *New Youth* was available, we define him as being exposed to Communist ideology. By this design, we study an intention-to-treat effect. In addition to this dummy exposure indicator, we present the results using a continuous variable that calculates the number of exposure years.

3 Descriptive Evidence and Research Design

Descriptive Evidence. Did exposure to the *New Youth*, more specifically to post-1919 *New Youth*, make a difference? We present two pieces of motivational evidence from the raw data to throw light on this question. The first is the pattern over time that suggests the relevance of the Communist ideology, created by dividing the individuals into an exposed group (those who was living in a prefecture with the *New Youth*) and a non-exposed group and examining the difference in probability of joining the CCP. As shown in Figure 3(a), this difference turned positive after 1919–20, which coincides with the timing the *New Youth*'s promotion of Communism. The pattern is very similar when we define *New Youth* exposure at the county level, as shown in 3(b).

The second piece is the pattern across space, which suggests a relevance of ideology exposure across localities. Here, we plot the numbers of CCP members among the group with post-1919 *New Youth* exposure across prefectures in Figure 4(a), and similarly, such numbers among the group without post-1919 *New Youth* exposure in Figure 4(b). On average, the mean ratios are 19.3% and 11%, indicating a higher CCP probability for the group with post-1919 *New Youth* exposure in the raw data.

Estimation Strategy. The main challenge with our research is the notion that post-1919 *New Youth* exposure may reflect other locality characteristics that could affect CCP membership. To address this, we control for locality fixed effects and focus on the variation built from circulation change of the *New Youth* and an individual's life events. We will also examine possible endogeneity concerns regarding each variation by studying different subsamples (e.g., by focusing on the localities always with the *New Youth*). Our baseline specification for prefecture-level exposure is as follows:

$$CCP_i = \beta NYExposure_{i,p,1919-22} + \alpha NYExposure_{i,p,1915-18} + Pref_p + X_i + \epsilon_i \quad (1)$$

where CCP_i is a dummy variable indicating whether individual i became a CCP member or not. $NYExposure_{i,p,1919-22}$ and $NYExposure_{i,p,1915-18}$ refer to the availability of the *New Youth* during 1915–1918 and 1919–1922 when i was located in prefecture $Pref_p$. $Pref_p$ is a set of prefecture fixed effects that controls for many locality characteristics. X_i includes the 13 variables on personal background. We also include 5-team fixed effects in some specifications in X_i to further account for their experience at Whampoa.

To allow for possible changes in different prefectures before and after 1919, we can further control for the interactions between prefecture fixed effects and the post-1919 dummy ($Pref_p \times Post1919$).

The cost of doing so results in only approximately half of the observations (358 individuals) still having variations. We present these results as additional evidence.

Our research also includes a study on county-level exposures. We compare individuals who lived in the same county and control for county fixed effects with county-level exposures. These county-level exposures provide more precise information at the cost of having smaller cells to compare.

Finally, we report the standard errors clustered at the home county level in our baseline estimations. In addition, we present standard errors two-way clustered at the prefecture and the year levels and implement a nonparametric permutation test that randomly assigns exposures and does not rely on parametric assumptions of error structure (Chetty, Looney and Kroft 2009).

Balance Tests. Our identification assumption is that, conditional on the localities an individual had been to, $NYExposure_{p,1919-22}$ is close to random. This assumption is reasonable to consider when an individual's relocation was primarily driven by economic opportunities and life events rather than the presence of the *New Youth*. To confirm, we conduct balance tests with respect to all considered personal background variables. In Columns (1)–(3) of Table 1, we divide individuals into three groups: (1) those who could access post-1919 *New Youth* in their county, (2) those who could access post-1919 *New Youth* in their prefecture but not their county, and (3) those who could not access post-1919 *New Youth* in their prefecture. Our analyses on prefecture-level exposures are built by comparing (1) and (2) against (3). As shown in Columns (4)–(5), the characteristics between those exposed to the post-1919 *New Youth* and those who weren't are not significantly different once we control for prefecture fixed effects. This would not be the case if we were to fail in considering these locality fixed effects.

Similarly, our analyses on county-level exposures compare (1) against (2) and (3). As shown in Columns (6)–(7), the characteristics between the exposed and everyone else are not significantly different once we control for county fixed effects. Again, this would not be true without county fixed effects. These results highlight the importance of controlling for locality fixed effects to account for the sorting of individuals.

In addition, once we consider locality size (measured by urban population size in 1920), there is no systematic correlation between exposure to the *New Youth* and the number of the Whampoa cadets in our sample before and after 1919 (Appendix A.5).

4 Results

4.1 Ideology Exposure and CCP Membership

Main Results. Exposure to post-1919 *New Youth* has a large impact on CCP membership. Columns (1)–(6) of Table 2 focus on prefecture-level exposure. After controlling for prefecture fixed effects in Column (1), the estimate for β in equation (1) is 0.109. By contrast, pre-1919 *New Youth* exposure had little impact on the CCP membership (Column (2)), highlighting the importance of ideological content. Column (3) compares the impacts of post- and pre-1919 *New Youth* and confirms the previous findings. Column (4) controls for the personal characteristics that are available for the entire sample (education background, birth year fixed effects, team fixed effects, and number of CCP references). Column (5) shows a similar estimate when restricting the sample to those in Teams 1–4 for whom we have more information. Column (6) further controls for additional personal characteristics such as family background, demographic information, and behavior traits. We use the coefficient (0.117) in Column (6) as our baseline estimate.

Columns (7)–(12) present the results for county-level exposure, following the same specifications in Columns (1)–(6) but replacing prefecture fixed effects with county fixed effects. The estimate for β increases to 0.172, reflecting a larger impact for a more precise exposure measure.

As shown in Table 3, our estimates are similar if we further control for $Pref_p \times Post1919$, indicating that our finding is robust with regard to considering changes in different prefectures before and after 1919. As previously mentioned, 358 individuals would still have variations after including this set of controls, explaining the increase in standard errors compared with our baseline estimates.

Among the personal characteristics, having a personal CCP reference in most of the specifications is associated with an increased chance of joining the CCP. This is consistent with research on the importance of social networks in political participation. In contrast, socioeconomic variables are unable to predict CCP membership. In the examination of the interaction effects between personal characteristics and post-1919 *New Youth* exposure (Appendix B.1), we do not find strong heterogeneities except for some complementarity between ideology exposure and social networks and also between ideology exposure and business family background. The relatively insignificant impact of personal characteristics is likely because the CCP faced a great deal of risk at this stage, and it is difficult to imagine what type of person might benefit more from joining such a revolution.

The impact of post-1919 *New Youth* exposure is robust to various ways of calculating standard

errors (Appendix B.2). Moreover, we conduct a nonparametric permutation test by randomly assigning the treated sample 10,000 times. As reported in Figure 5, our estimate on post-1919 *New Youth* is significantly different from those associated with random assignments (with a p -value smaller than 0.001). Since these results do not rely on parametric assumptions, they further support the relevance of post-1919 *New Youth* exposure.

Moreover, we employ a continuous measure for $NYExposure_{p,1919-22}$ by counting the share of years exposed (instead of a dummy variable) and find a pattern similar to using a dummy variable (Appendix B.3). For the following analyses, we focus on the dummy variable of exposure to facilitate interpretations.

A Closer Look at Possible Endogeneity Concerns. While we have considered time-invariant locality characteristics and possible changes in the influences of those locality characteristics before and after 1919 in our main analyses, here, we discuss issues revolving around (1) the possible endogenous entry of the *New Youth* and (2) possible relocation preferences with regard to those interested in reading the *New Youth*.

The finding that $NYExposure_{p,1915-18}$ has little impact on an individual's choice to join CCP choice suggests that endogenous entry of the *New Youth* is not a critical concern (Table 2). As a further check, we separate the localities that always had the *New Youth* and those that experienced the entry and exit of the *New Youth*. As shown in Table 4, the estimates are similar in either case.

To check whether an individual's relocation prefecture could be a concern, we construct placebo relocations by assuming an individual's location were in the previous one or the next one in his life history. We use these placebo relocations to calculate placebo exposures to post-1919 *New Youth*. In principle, these placebo exposures capture one's potential relocation preferences. As shown in Table 5, these placebo exposures cannot explain our finding on $NYExposure_{p,1919-22}$, thus, our main finding is also unlikely to be explained by an individual's relocation in response to the availability of post-1919 *New Youth*.

Finally, we separate our comparison group into two subgroups: (i) those without access to post-1919 *New Youth* and (ii) those who had been in localities with post-1919 *New Youth* available but missed the magazine. Group (ii)'s relocation is not to be driven by the presence of post-1919 *New Youth*, as the magazine was available but these individuals missed it. Table 6 presents the estimate created by comparing the exposed group to group (ii). This estimate (0.12) is close to our baseline ones, demonstrating yet again that our finding is not driven by the sorting of individuals

toward post-1919 *New Youth*.

Measurement Error. Measurement error is likely. We conduct three sets of analyses at the locality and individual levels in Appendix B.4. First, our estimate is robust to dropping major metropolitan cities (Beijing, Shanghai, Guangzhou), implying that our findings are not driven by just a few large cities. In our following analysis on persuasion rate, we are further able to examine the heterogeneity across localities.

Our second and third examinations concern the precision of life events in an individual's records. The second shows that our finding is not dependent on the number of life events and robust to dropping those with few or many life events recorded. Thus, our finding is not driven by either the abundance or a lack of records on individuals.

Third, as previously mentioned, we infer certain missing years based on an individual's other records with concrete year information and formulate a calculation against the timeline of the education system in this era. To check whether our method is reasonable, we apply the same inference method by pretending that we only knew the birth year, end year, and life events for the individuals in this subsample. As shown in Appendix B.4, the estimate is similar to our baseline if we apply the same inference method to all our sample, which solidifies the validity of inferring missing years on an individual's records.

Gauging the Persuasion Rate. To better understand the magnitude of our estimates, we would like to calculate the persuasion rate following DellaVigna and Gentzkow (2010), where the persuasion rate (P) is defined as follows:

$$P = \frac{y_T - y_C}{e_T - e_C} \frac{1}{1 - y_0} \quad (2)$$

While we know $y_T - y_C$ (i.e., the impact of being exposed) based on our estimates and can proxy y_0 by y_C (i.e., the probability without being exposed), we do not know the reading rate difference ($e_T - e_C$). The difference between the impacts of prefecture and county exposures provides some useful benchmark to gauge $e_T - e_C$. Specifically, it is reasonable to assume that the reading probability is close to zero if the *New Youth* is available outside a county (i.e., $e_C = 0$). The logic is as follows: Without considering the reading rates, the prefecture-level P is $0.117/(1-0.122)=13.3\%$ and the county-level P is $0.172/(1-0.092)=18.9\%$, which implies that the reading probability at the prefecture level should be 70.8% ($13.3/18.9$) at the county level. This probability is close to that from assuming a zero reading probability if the *New Youth* is available outside a county:

$365/(365+143)=71.9\%$.

Our inferred e_C is confirmed in studying whether exposure to post-1919 *New Youth* in a given county affects a cadet's choice to join the CCP when located in other counties in the same prefecture. In this exercise, we find no significant spillover effect (Columns (1)–(2) of Appendix B.5).

To gauge the magnitude of e_T , we exploit two heterogeneous patterns in circulation locality. First, we separate the circulation county into small and big cities, as defined by a threshold of urban population size of 100,000. Using the same specification as in our baseline, we obtain estimates of 0.26 for small cities and 0.13 for big cities (Columns (3)–(4) of Appendix B.5). Second, a subsample of the circulation points were bookstores located in schools. Based on the life history of the cadets, we know whether a cadet was studying in those schools. If the *New Youth* was available at a cadet's school, it seems reasonable to assume e_C being close to 1 given that the *New Youth* was the most discussed periodical in this era. When separating this very local exposure from others, we find its coefficient to be 0.26 whereas that of the rest is 0.15 (Columns (5)–(6)). Taking 0.26 as the upper bound, we obtain a persuasion rate of 29.2%. The relatively high persuasion rate is likely because we are studying a group of radical youth.

4.2 Interpreting Post-1919 *New Youth* Exposure

Motivated by historical narratives (Section 2.1), we use the post-1919 *New Youth* exposure to measure Communist ideology exposure. This is not meant to deny the other periodicals these individuals likely read or that various activist organizations emerging in the New Culture Movement contributed in their own way to the recruitment of the CCP. Here, we are examining whether post-1919 *New Youth* could serve as a major channel for Communist ideology diffusion. To this end, we use archival information on pro-Communism periodicals and activist groups in this era to understand the importance of the *New Youth* as a channel for Communist ideology diffusion.

The Network of Pro-Communism Periodicals. The *New Youth* inspired other pro-Communism periodicals to start, with the editors of the *New Youth* helping to co-found several of those periodicals. Based on a series of archives published as *Red Collections: Progressive Periodicals from 1915–1949* (Xue 2014), we gather information on the editors and editorial offices for all 19 pro-Communism periodicals that existed during 1919–1922 in these archives. We define a link between two periodicals as sharing an editor or editorial office.

We plot the network among these pro-Communism periodicals in Figure 6. As shown by the closeness centrality index noted in the figure, the *New Youth* is indeed the most central periodical in the network. In comparison to the *New Youth*, the other periodicals started later and thus had a shorter period of influence during 1919–1922.

We also gather information on the circulation locations of two other pro-Communism periodicals in the network, *Young China* and the *Guide Weekly*. We specifically chose these because of their closeness in circumstance to the *New Youth*. *Young China* had the same length of existence as the post-1919 *New Youth* during 1919–1922 and the *Guide Weekly* is the second central periodical in the network. We find that exposures to these additional periodicals can be considered a subset of exposure to the post-1919 *New Youth*: 47% of our studied cadets were exposed to the circulation of these two periodicals, out of which 99.4% overlapped with post-1919 *New Youth* exposure. When separating individuals out between those exposed to post-1919 *New Youth* only and those exposed to all three periodicals, we find similar estimates (Panel (a) of Table 7), supporting the importance of post-1919 *New Youth* as a reasonable proxy for Communism diffusion.

Activist Groups and Communist Organizations. There existed various voluntary groups in the New Cultural Movement. We focus on the 28 of them that have been characterized as left-wing activist groups from archival information of their reports and meeting minutes (Zhang, Yin, Hong and Wang 1979, Xu 2013).²¹ To check the relationship between these organizations and ideology exposure, we code an exposure measure to these organizations using the same way we code exposure to the *New Youth*. When examining the relationship between these exposures, we find a heavy overlapping pattern: 31% of our studied individuals were exposed to those activist groups, out of which 99.1% overlapped with post-1919 *New Youth* exposure. When separating out the individuals that were only exposed to post-1919 *New Youth* and those exposed to both post-1919 *New Youth* and these 28 activist groups, we find an estimate of 0.11 for the former and an estimate of 0.18 for the latter (Panel (b) of Table 7). These results suggest complementarity between ideas and social groups. Nevertheless, exposure to post-1919 *New Youth* has a major impact in localities even without these activist groups. One interpretation is that post-1919 *New Youth* provides a venue for intellectuals to share Communism ideas with a broad audience across the nation.

In sum, we interpret the role of post-1919 *New Youth* as a reasonable proxy for Communism

²¹Zhang, Yin, Hong and Wang (1979) provides the data source. Xu (2013) studies why five of these 28 groups became Communist groups. We follow Xu (2013) in identifying these groups.

diffusion. While it is possible that our studied individuals were exposed to additional periodicals and organizations, considering their impact does not appear to change our main finding.

4.3 Future Political Outcomes

Do individuals mobilized by ideas behave differently in critical historical moments? We leverage information on the future careers of the Whampoa cadets to shed light on this important yet difficult question. During the 1920s–1940s, the CCP and the KMT experienced various political struggles. In this process, many quit their parties and sacrificed their lives. For instance, in 1927, the KMT under Chiang Kai-shek’s direction worked out a plan to “clean the party” of Communist influence, resulting in a purge in which the CCP lost over 90% of its members. The two civil wars (1927–1936 and 1945–1949) and the anti-Japanese war (1937–1945) were also significant conflicts that left both parties experiencing a loss of members.

We compare four groups of individuals: CCP members with and without post-1919 *New Youth* exposure and KMT with and without post-1919 *New Youth* exposure. Table 8 presents comparisons using KMT without post-1919 *New Youth* as the reference group, where we use seemingly unrelated regressions to allow for correlated error terms across the equations (for sacrifice and quitting). As shown, the shares of quitting or sacrifice are higher for the CCP than those for the KMT, reflecting a greater uncertainty faced by the CCP. Moreover, within the CCP, there appears to be a striking difference: Those who were exposed to post-1919 *New Youth* were 23 percentage points less likely to quit the party and 25 percentage points more likely to sacrifice their lives (Columns (1)–(4)).

As we take a closer look at different periods, we find that our finding above is primarily driven by the period from 1927 to 1936, during which the CCP was weaker and was under severe attacks by the KMT. Out of the 66 CCP members who sacrificed their lives, 27 died during 1924–27 and 33 died during the first KMT-CCP Civil War. Out of the 38 CCP members who quit the party, 5 quit during 1924–27, and 33 quit during the first KMT-CCP civil war. A similar pattern emerges when zooming into the period of the first civil war (Columns (5)–(6)). These results provide suggestive evidence that individuals mobilized by ideas can behave differently, especially in tumultuous times.

Individual choices at critical historical moments are, of course, likely to be shaped by a variety of factors. As emphasized by historians, most revolutionaries were gradually transformed to be committed Communists via the complicated process of revolution (Van de Ven 1992, Esherick 1995). Therefore, our finding does not imply that future political outcomes are predetermined, but

rather suggests that ideology exposure can interact with other factors to influence personal choices.

5 Conclusion

Ideas and their diffusion may have played pivotal roles in many social and political movements. One may even claim: “no ideas, no revolution” (Brinton 1938).²² For instance, Bailyn (1967) emphasizes that the American Revolution had an ideological grounding centered on the struggle between power and liberty, which is further argued to have shaped the American identity (Huntington 1981). Likewise, the French Revolution was motivated by a set of ideas encapsulated as liberty, equality, and fraternity. Such arguments, however, can be contested due to the challenges of evaluating the specific impact of ideas. In this study, we focus on the ideas of Communism, which have had profound and global impacts yet have not been studied significantly by social scientists. We measure Communism exposure by combining information on the supply of ideas with the biographical life history information on a group of key individuals. Through the process, we are able to demonstrate the importance of ideology exposure in shaping the political choice of a group of individuals who would influence the fate of modern China.

To be certain, ideology exposure is not ever going to be the only determinant of one’s political choice. On the contrary, we find that the political identity of one’s personal references also serves as another strong predictor of an individual’s own political identity, suggesting the strength of the influence that social networks can have. At the same time, the association between one’s socioeconomic background and political identity is found to be minimal in our specific context. The particular relevance of ideology, in contrast with socioeconomic status, reflects a general uncertainty faced by the CCP during our studied period. These findings are likely to be more relevant for movements that face high risk. While it is generally difficult to study individual participation in high-risk social movements, our study offers a path to depict the profile of individual participants and, in turn, understand their political choices. We believe that this approach can be applied to other settings where systematic information exists on a selected group of relevant individuals.

²²As noted by Brinton, this does not mean that ideas cause revolution. The processes of ideas and revolution are often mutually dependent, which creates challenges in identifying the influence of ideas.

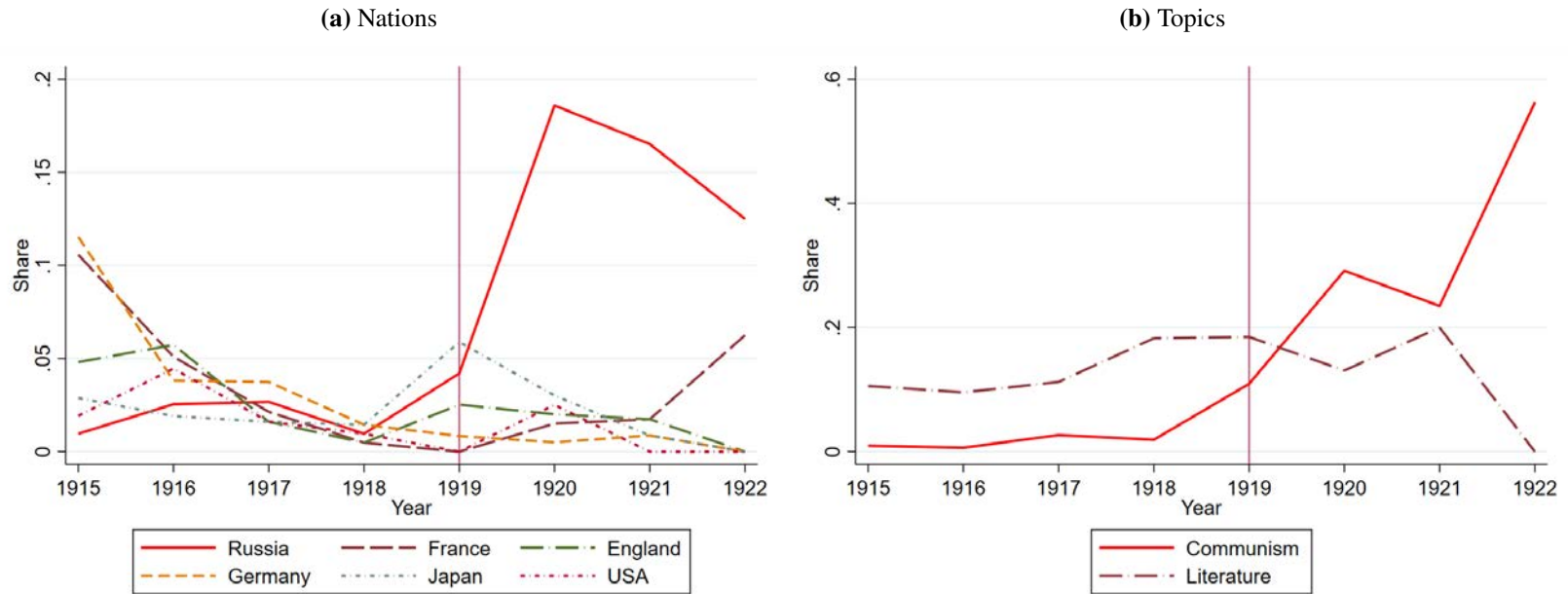
References

- [1] Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya (2015), “Radio and the rise of the Nazis in Prewar Germany,” *Quarterly Journal of Economics*, 130(4): 1885–1939.
- [2] Ash, Elliott, Daniel L. Chen, and Suresh Naidu (2022), “Ideas have consequences: The impact of law and economics on American justice,” NBER Working Paper 29788.
- [3] Bai, Ying, Ruixue Jia, and Jiaojiao Yang (2022), “Web of power: How elite networks shaped war and politics in China,” *Quarterly Journal of Economics*, forthcoming.
- [4] Becker, Sascha O., and Luigi Pascali (2019), “Religion, division of labor, and conflict: Anti-Semitism in Germany over 600 years,” *American Economic Review*, 109(5): 1764–1804.
- [5] Bailyn, Bernard (1967), *The Ideological Origins of the American Revolution*. Harvard University Press.
- [6] Berlin, Isaiah (1979), *Against the Current: Essays in the History of Ideas*. London: Pimlico.
- [7] Brinton, Crane (1938), *The Anatomy of Revolution*. New York: Vintage Books. Reprinted in 1965.
- [8] Cagé, Julia, Anna Dagorret, Pauline A. Grosjean, and Saumitra Jha (2022), “Heroes and villains: The effects of combat heroism on autocratic values and Nazi collaboration in France,” *American Economic Review*, forthcoming.
- [9] Cantoni, Davide (2012), “Adopting a new religion: the case of Protestantism in 16th century Germany,” *Economic Journal*, 122(560): 502–531.
- [10] Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang (2017), “Curriculum and ideology,” *Journal of Political Economy*, 125(2): 338–392.
- [11] Central Executive Committee of the Kuomintang (1924), *Lujun junguan xuexiao xiangxi diaochabiao* [Detailed Survey for the First to Fourth Teams of the First Cohort of the Military Academy]. reprinted in *Jindai zhongguo shiliao congkan*, edited by Yunlong Shen, 1990. Taipei: Wenhai Press.
- [12] Chen, Duxiu (1916), “Wuren Zuihou Zhi Juewu [Our Final Awakening],” *New Youth*, Vol.1(6).
- [13] Chen, Ting, and James Kai-Sing Kung (2020), “The rise of Communism in China,” Working Paper.
- [14] Chen, Yuhuan (2017), *Whampoa Junxiao Diyiqisheng Quanjilu* [Complete Records on the First-Cohort Cadets of the Whampoa Military School]. Sixing Wenhua Press: Taipei.
- [15] Chetty, Raj, Adam Looney, and Kory Kroft (2009), “Salience and taxation: Theory and evidence,” *American Economic Review*, 99(4): 1145–77.
- [16] Chow, Tse-tsung (1960), *The May Fourth Movement: Intellectual Revolution in Modern China*. Harvard University Press.
- [17] DellaVigna, Stefano, and Matthew Gentzkow (2010), “Persuasion: empirical evidence,” *Annual Review of Economics*, 2(1): 643–669.
- [18] DellaVigna, Stefano, and Ethan Kaplan (2007), “The Fox News effect: Media bias and voting,” *Quarterly Journal of Economics*, 3(122): 1187–1234.
- [19] Dirlik, Arif (1989), *The Origins of Chinese Communism*. Oxford University Press.
- [20] Dippel, Christian, and Stephan Heblich (2021), “Leadership in social networks: Evidence from the Forty-Eighters in the Civil War,” *American Economic Review*, 111(2): 472–505
- [21] Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2020), “Social Media and Protest Participation: Evidence from Russia,” *Econometrica*, 88(4): 1479–1514.
- [22] Esherick, Joseph (1995), “Ten theses on the Chinese revolution,” *Modern China*, 21(1): 45–76.

- [23] Fan, Guorui, and Thomas S. Popkewitz eds. (2020), *Handbook of Education Policy Studies*. Vol.2. Springer Nature.
- [24] Graff, David Andrew, and Robin Higham, eds. (2012), *A Military History of China*. University Press of Kentucky.
- [25] Hirschman, Albert O. (1977), *The Passions and the Interests: Political Arguments for Capitalism Before Its Triumph*. Princeton: Princeton University Press, 1977.
- [26] Huntington, Samuel P. (1981), *American Politics: The Promise of Disharmony*. Harvard University Press.
- [27] Jin, Guantao and Qingfen Liu (2008), *Guannianshi yanjiu: Zhongguo xiandai zhongyao zhengzhi shuyu de xingcheng [Studies in the History of Ideas: The Formation of Important Modern Chinese Political Terms]*. Hong Kong: Chinese University of Hong Kong Press.
- [28] Johnson, Chalmers A (1962), *Peasant Nationalism and Communist Power: the Emergence of Revolutionary China, 1937-1945*. Stanford University Press.
- [29] Sung-K'ang, Huang (1964), *Li Ta-chao and the Impact of Marxism on Modern Chinese Thinking*. Paris: Mouton & Co.
- [30] Ishikawa, Yoshihiro (2013), *The Formation of the Chinese Communist Party*. Columbia University Press.
- [31] Jones, Benjamin F., and Benjamin A. Olken (2005), "Do leaders matter? National leadership and growth since World War II," *Quarterly Journal of Economics*, 120(3): 835–864.
- [32] Magness, Phillip, and Michael Makovi (2002), "The Mainstreaming of Marx: Measuring the effect of the Russian Revolution on Karl Marx's influence," *Journal of Political Economy*, forthcoming.
- [33] Meisner, Maurice (1999), *Mao's China and After: A History of the People's Republic*. Simon and Schuster.
- [34] Michalopoulos, Stelios, Alireza Naghavi, and Giovanni Prarolo (2018), "Trade and geography in the spread of Islam," *Economic Journal*, 128(616): 3210–3241.
- [35] Mokyr, Joel (2005), "The intellectual origins of modern economic growth," *Journal of Economic History*, 65(2): 285–351.
- [36] Prat, Andrea, and David Strömberg (2013), "The political economy of mass media," in *Advances in economics and econometrics* Vol. 2.
- [37] Satyanath, Shanker, Nico Voigtlönder, and Hans-Joachim Voth (2017), "Bowling for Fascism: Social Capital and the Rise of the Nazi Party," *Journal of Political Economy*, 125(2): 478–526.
- [38] Schwarcz, Vera (1986), *The Chinese Enlightenment: Intellectuals and the Legacy of the May Fourth Movement of 1919*. University of California Press.
- [39] Smith, Stephen Anthony (2000), *A Road is Made: Communism in Shanghai, 1920–27*. Honolulu: University of Hawaii Press.
- [40] Smith, Stephen Anthony, ed. (2014), *The Oxford Handbook of the History of Communism*. Oxford.
- [41] Snow, Edgar (1937), *Red Star Over China: The Classic Account of the Birth of Chinese Communism*.
- [42] Taylor, Jay (2009), *The Generalissimo: Chiang Kai-shek and the Struggle for Modern China*. Harvard University Press.
- [43] Van de Ven, Hans (1992), *From Friend to Comrade: The Founding of the Chinese Communist Party, 1920–1927*. University of California Press.
- [44] Xu, Xiaohong (2013), "Belonging before believing: Group ethos and bloc recruitment in the making of Chinese communism," *American Sociological Review*, 78(5): 773–796.
- [45] Xue, Muqiao (ed. 2014). *Hongcang: Jinbu Qikan Zonghui (1915–1949) [Red Collections Progressive Periodicals from 1915–1949]*. Xiangtan: Xiangtan University Press.

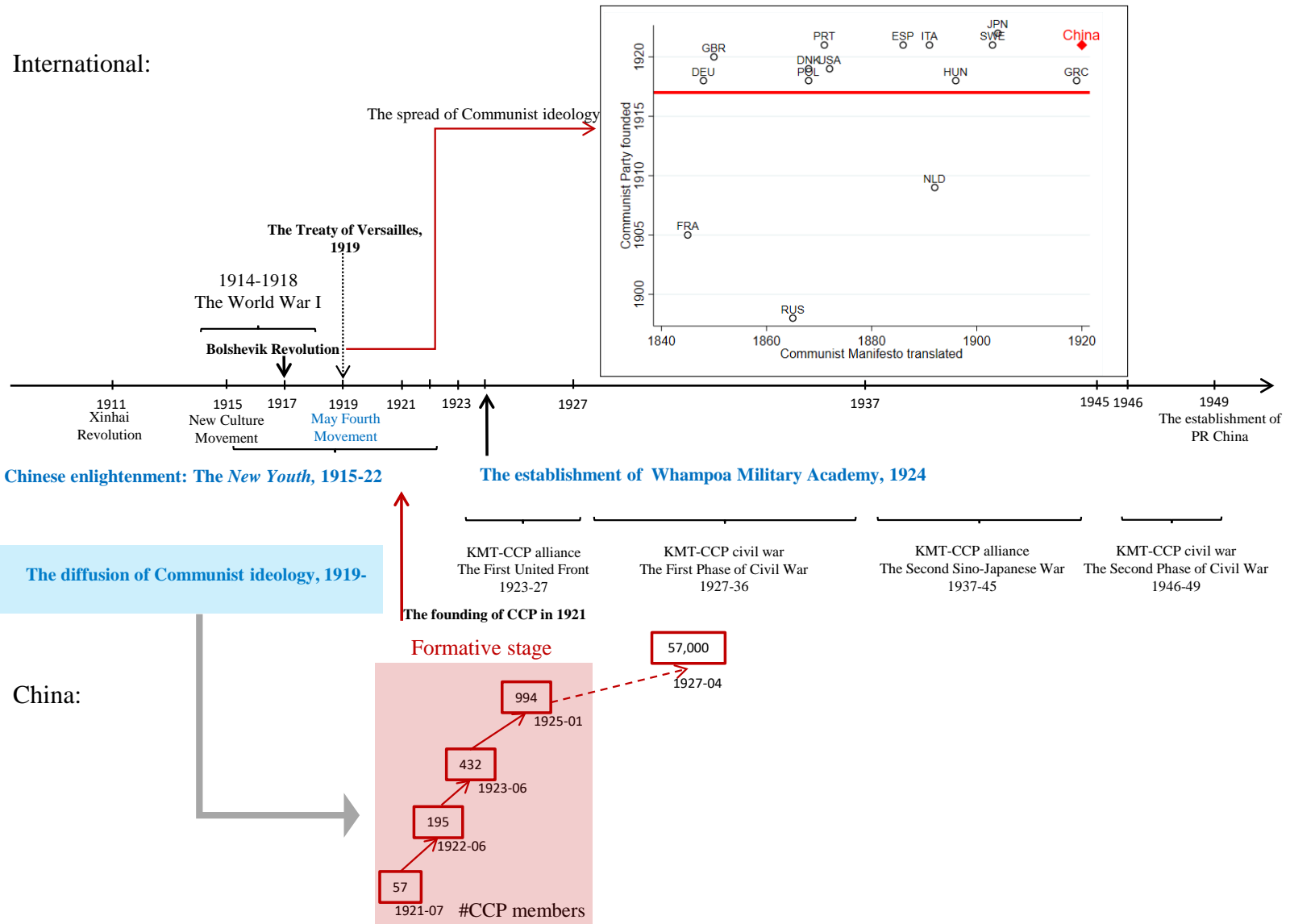
- [46] Yang, Kuisong (1994), “Mosike xinfaxian de Deng Xiaoping zizhuan [“Autobiography of Deng Xiaoping” newly discovered in Moscow],” *Ming Pao Monthly*, Vol. 9.
- [47] Yeh, Wen-hsin (1996), *Provincial Passages: Culture, Space, and the Origins of Chinese Communism*. University of California Press.
- [48] Yun, Daiying (1981), *Yun Daiying riji [Diaries of Yun Daiying]*. Beijing: Central Party School Press.
- [49] Zhang Yunhou, Xuli Yin, Qingxiang Hong, and Yunkai Wang (ed. 1979), *Wusi shiqi de shetuan [Voluntary Societies during the May Fourth Era]*. Beijing: Sanlian chubanshe.
- [50] Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020), “Political effects of the internet and social media,” *Annual Review of Economics*, Vol. 12: 415–438.

Figure 1. Ideological Transition of the *New Youth*



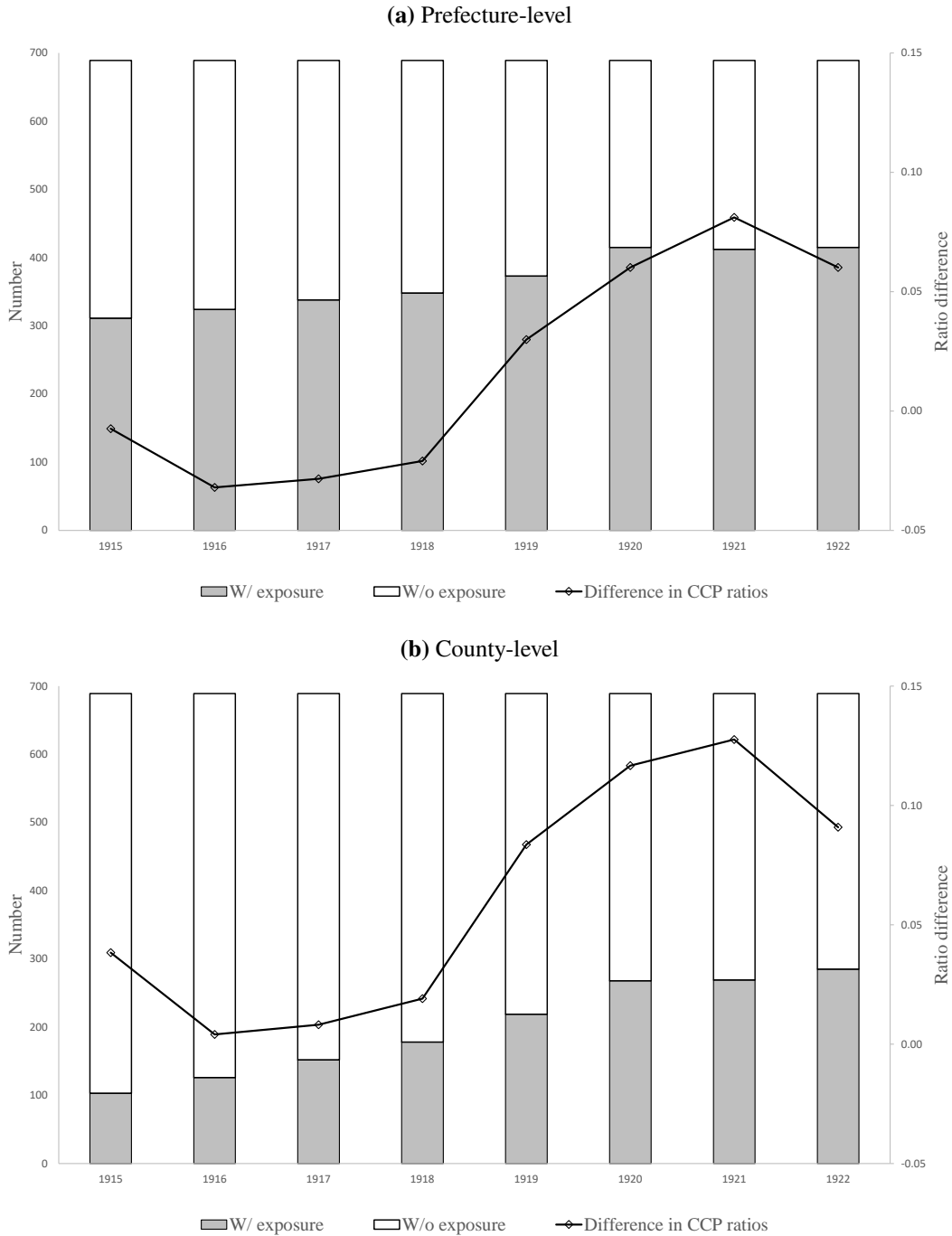
Note. This figure shows the ideological transition of the *New Youth* around 1919–20. In Panel (b), the "Communism" topic includes the phrases "Marx", "Lenin", "socialism", "communism", "Soviet", "Bolshevism", "revolution", "labor", "worker", "strike", and "proletariat". "Literature" includes all novels, dramas, literary critics, essays on linguistics and philology, and book-reading notes. We exclude poems when calculating these shares.

Figure 2. Key Events (1911–1949)



Note. Our analyses focus on how exposure to the *New Youth* (1919–22) affected (1) the Whampoa cadets’ choice to become a CCP member in CCP’s formative stage (1921–27) and (2) the outcomes of these cadets during 1920s-49.

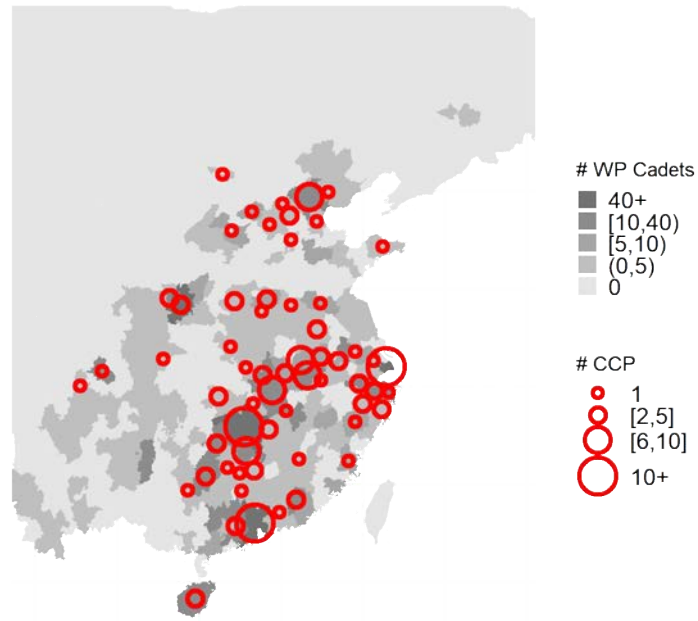
Figure 3. Raw Data I: Exposure to the *New Youth* and CCP Membership



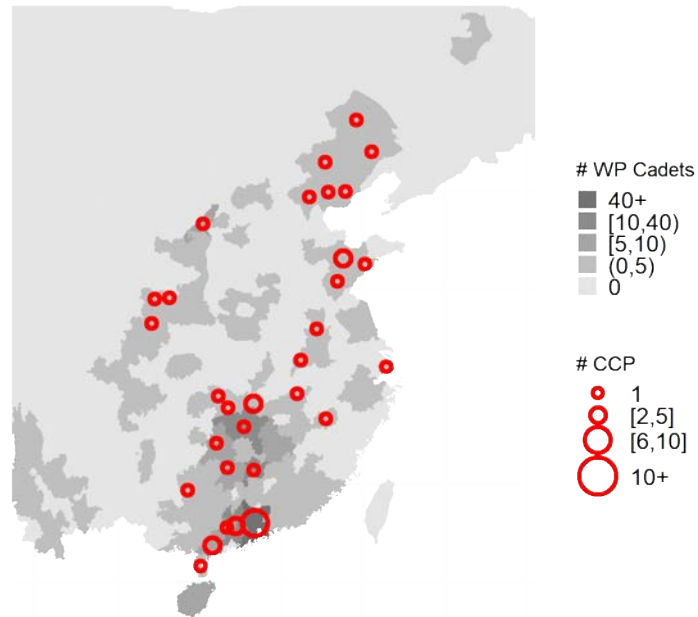
Note. The shaded and unshaded areas in the bar represent the number of cadets with and without *New Youth* exposure in each year respectively. The solid line with diamond markers represents the difference in CCP ratios between the cadets with *New Youth* exposure and those without. As shown, the differences became positive after 1919.

Figure 4. Raw Data II: Post-19 *New Youth* Exposure and CCP Membership

(a) CCP w/ Post-19 *New Youth* Exposure



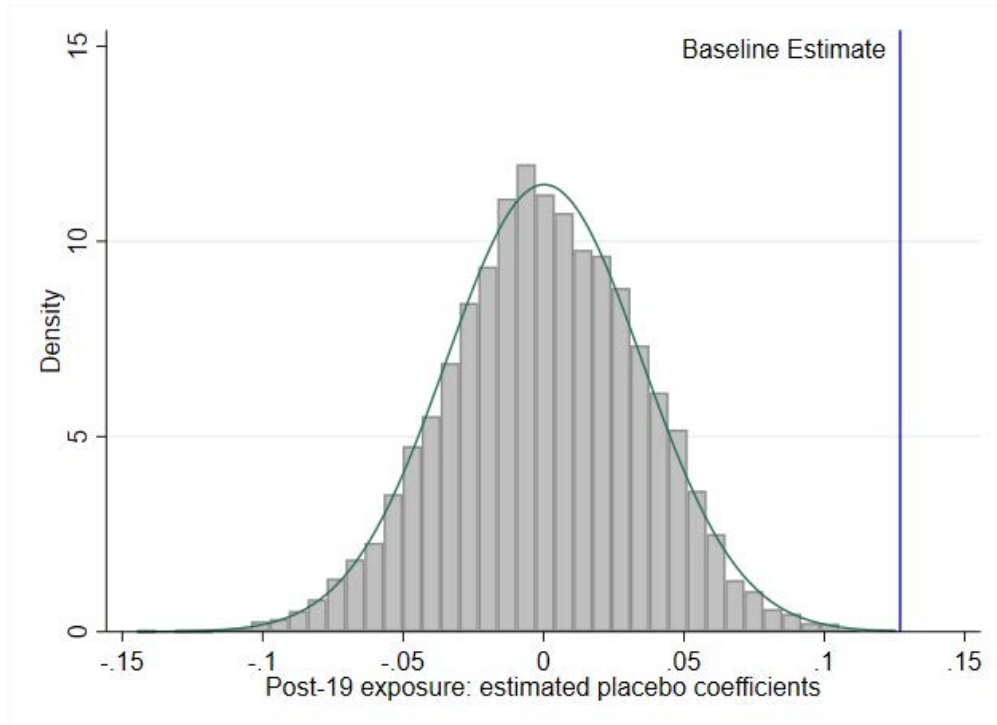
(b) CCP w/o Post-19 *New Youth* Exposure



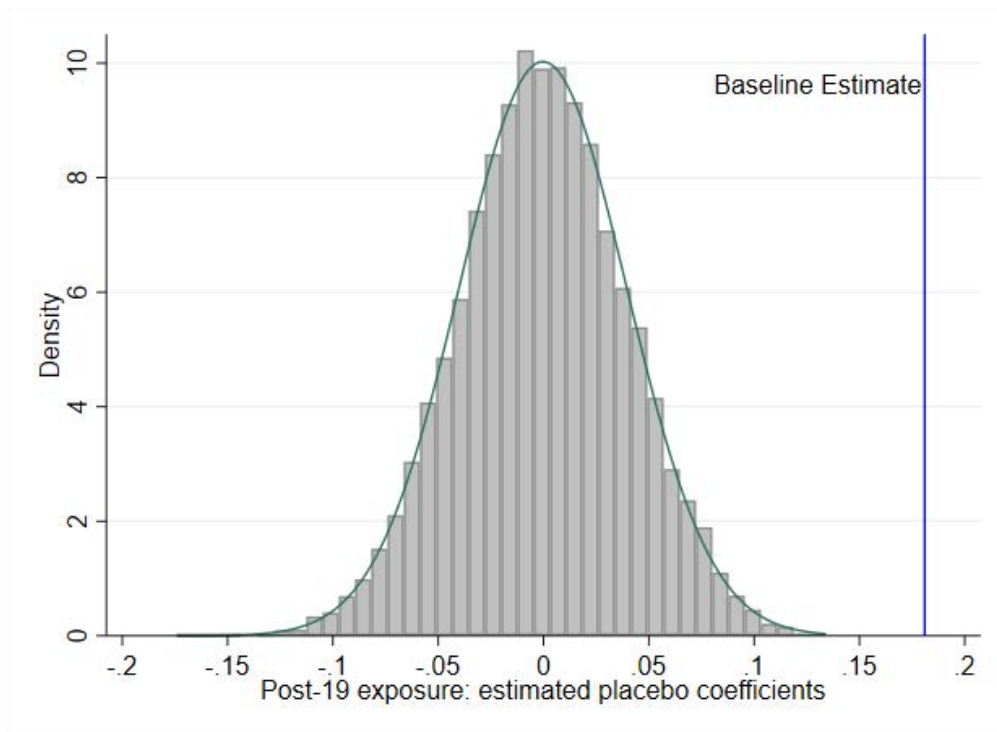
Note. These figures plot the number of CCP members and the number of cadets across the groups with or without post-19 *New Youth* exposure.

Figure 5. Distribution of Placebo Estimates

(a) Exposure Defined at Prefecture Level

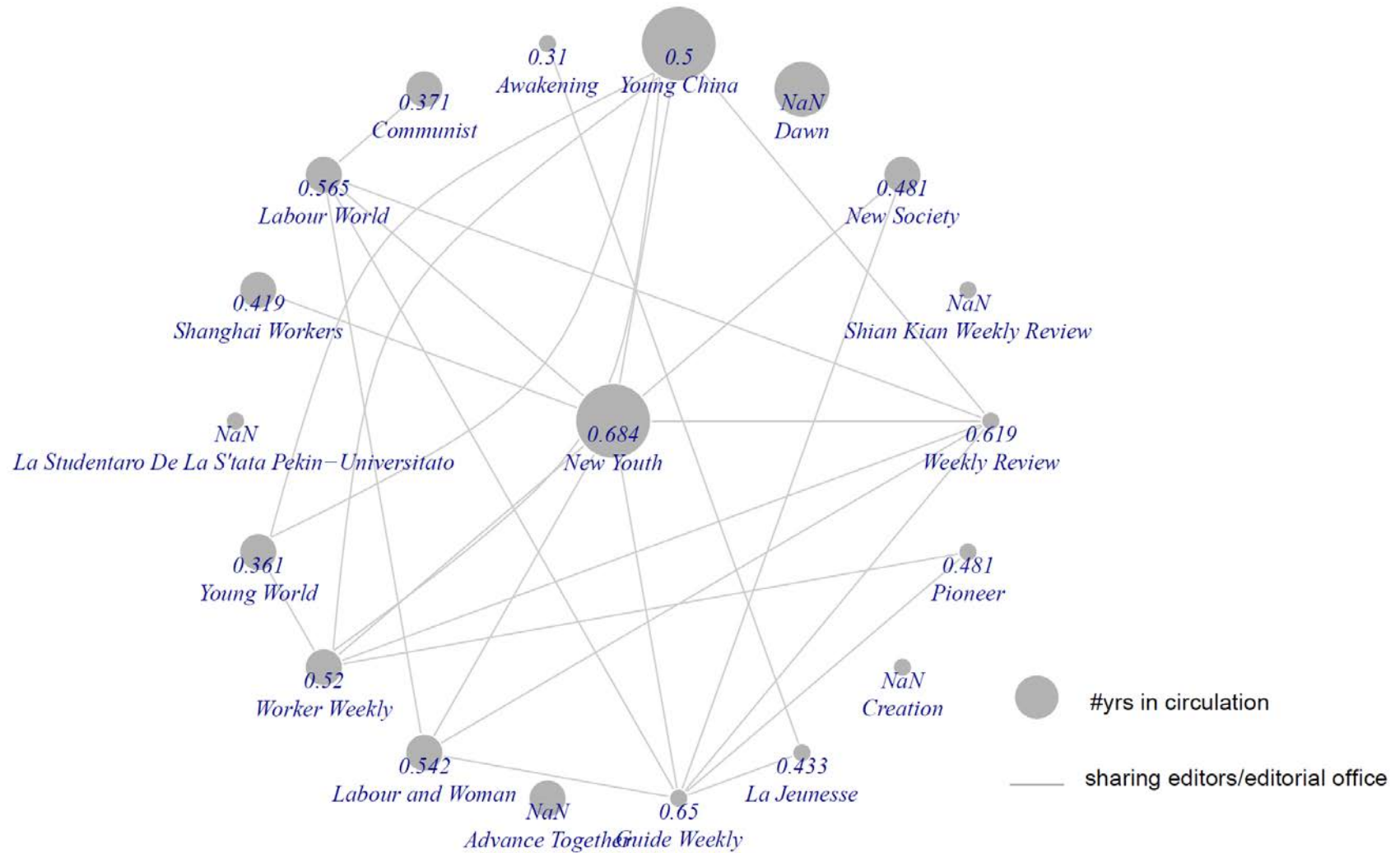


(b) Exposure Defined at County Level



Note. These figures plot estimates from randomly assigned exposures. They demonstrate that our estimates are significantly different from these randomly assigned treatment effects.

Figure 6. The Network of Pro-Communism Periodicals



Note. This figure plots links among all 19 Pro-Communism periodicals in existence during 1919–1922. A link means that two periodicals share editors or an editorial office. The size of the circles indicates the number of circulation years during 1919–1922, and the number in the circle indicates the closeness centrality of a periodical in this network.

Table 1. Balance Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	New Youth exposure, 1919-22			Differences			
Prefecture-level exposure	=1	=0	=0	(1)+(2) vs. (3)		(1) vs. (2)+(3)	
County-level exposure	=1	=0		w/o Pref. FE	w/ Pref. FE	w/o county FE	w/ county FE
Sample: All							
Personal characteristics I, N=	365	143	181				
Birth year	1900.5 (2.889)	1901.2 (2.802)	1900.7 (2.911)	0.009 [0.260]	0.297 [0.457]	-0.445** [0.218]	0.223 [0.491]
Education: Middle school and above	0.981 (0.137)	0.930 (0.256)	0.928 (0.259)	0.038* [0.023]	0.012 [0.037]	0.052*** [0.016]	0.001 [0.042]
Education: vocational	0.633 (0.483)	0.615 (0.488)	0.541 (0.500)	0.087 [0.055]	-0.055 [0.065]	0.059 [0.051]	-0.020 [0.108]
Education: college	0.164 (0.371)	0.035 (0.184)	0.050 (0.218)	0.078*** [0.025]	0.000 [0.035]	0.121*** [0.024]	-0.032 [0.050]
Networks: # CCP references	0.263 (0.608)	0.049 (0.247)	0.116 (0.570)	0.087* [0.051]	0.069 [0.084]	0.177*** [0.042]	0.090 [0.113]
Sample: Teams 1-4:							
Personal characteristics II, N=	314	85	131				
Family: gentry	0.280 (0.450)	0.188 (0.393)	0.214 (0.412)	0.047 [0.043]	0.042 [0.082]	0.077** [0.039]	0.132 [0.118]
Family: business	0.303 (0.460)	0.294 (0.458)	0.328 (0.471)	-0.027 [0.049]	0.007 [0.082]	-0.012 [0.041]	-0.058 [0.139]
Family economic status (1-3)	1.675 (0.521)	1.518 (0.569)	1.611 (0.549)	0.031 [0.060]	0.102 [0.098]	0.101** [0.050]	0.051 [0.148]
# Siblings	3.274 (2.258)	3.624 (2.502)	3.489 (2.761)	-0.140 [0.254]	0.182 [0.558]	-0.268 [0.216]	-0.292 [0.879]
Married	0.506 (0.501)	0.471 (0.502)	0.679 (0.469)	-0.181*** [0.051]	-0.141 [0.092]	-0.091* [0.048]	0.005 [0.146]
Having children	0.146 (0.354)	0.129 (0.338)	0.122 (0.329)	0.021 [0.035]	0.001 [0.070]	0.021 [0.031]	0.073 [0.102]
Religious	0.099 (0.299)	0.129 (0.338)	0.122 (0.329)	-0.017 [0.035]	0.045 [0.047]	-0.026 [0.029]	-0.006 [0.073]
Smoking and drinking	0.035 (0.184)	0.106 (0.310)	0.053 (0.226)	-0.003 [0.021]	0.028 [0.032]	-0.039* [0.022]	-0.100 [0.066]

Note. Standard deviations are reported in the parentheses in Columns (1)–(3). Standard errors clustered at the home county level are presented in brackets in Columns (4)–(7). Once locality fixed effects are included, the groups with or without exposure are balanced across all characteristics (Columns (5) and (7)).

Table 2. Baseline Estimates: Impact of Post-19 *NewYouth* Exposure on CCP Membership
Dependent Variable: CCP=1/0

New Youth Exposure defined at	Prefecture level						County level					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Teams 1-4						Teams 1-4	
New Youth Exposure, 1919-22	0.109** (0.048)		0.128** (0.050)	0.127** (0.052)	0.127** (0.058)	0.117** (0.056)	0.153** (0.077)		0.152** (0.077)	0.181** (0.073)	0.162** (0.081)	0.172** (0.082)
New Youth Exposure, 1915-18		-0.047 (0.046)	-0.081* (0.048)	-0.055 (0.051)	-0.091 (0.066)	-0.094 (0.068)		-0.044 (0.099)	-0.038 (0.097)	-0.006 (0.095)	-0.103 (0.115)	-0.104 (0.116)
Education: vocational				0.009 (0.046)	0.022 (0.047)	0.018 (0.047)				-0.066 (0.067)	-0.074 (0.075)	-0.078 (0.076)
Education: college				-0.028 (0.062)	-0.003 (0.073)	-0.002 (0.075)				0.041 (0.138)	0.003 (0.159)	-0.005 (0.179)
#CCP references				0.181*** (0.048)	0.171*** (0.050)	0.173*** (0.049)				0.179** (0.078)	0.076 (0.089)	0.079 (0.092)
Family: gentry						-0.015 (0.052)						-0.030 (0.092)
Family: business						0.021 (0.052)						0.004 (0.090)
Family economic status (1-3)						0.022 (0.039)						-0.011 (0.059)
#Siblings						-0.001 (0.007)						-0.000 (0.011)
Married						-0.072* (0.043)						-0.022 (0.078)
Having children						-0.012 (0.064)						-0.076 (0.101)
Religious						-0.030 (0.051)						-0.069 (0.090)
Smoking and drinking						0.042 (0.087)						0.042 (0.143)
Prefecture FE	Y	Y	Y	Y	Y	Y						
County FE							Y	Y	Y	Y	Y	Y
Team FE				Y	Y	Y				Y	Y	Y
Birth year FE				Y	Y	Y				Y	Y	Y
Observations	689	689	689	689	530	530	689	689	689	689	530	530
R-squared	0.386	0.381	0.389	0.443	0.534	0.542	0.560	0.552	0.560	0.603	0.747	0.750

Note. Columns (1)–(4) and (7)–(10) include all cadets whereas Columns (5)–(6) and (11)–(12) focus on Team 1–4 on whom we have more information. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1

Table 3. Controlling for Prefecture FEs \times Post 1919
Dependent Variable: CCP=1/0

New Youth Exposure defined at	Prefecture level		County level	
	(1)	(2)	(3)	(4)
New Youth Exposure, 1919-22	0.126** (0.054)	0.125** (0.057)	0.150* (0.085)	0.181** (0.083)
New Youth Exposure, 1915-18	-0.081 (0.050)	-0.053 (0.053) (0.047) (0.063) (0.049)	-0.043 (0.101)	-0.006 (0.100) (0.072) (0.144) (0.090)
Prefecture FE	Y	Y		
County FE			Y	Y
Prefecture FE * 1919-22	Y	Y	Y	Y
Personal characteristics I	Y	Y	Y	Y
Team FE		Y		Y
Observations	689	689	689	689
R-squared	0.391	0.445	0.569	0.611

Note. After controlling for Pref FE*Post, the variation stems from 358 individuals. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1

Table 4. Always-New Youth vs. Not-always-New Youth
Dependent Variable: CCP=1/0

	(1)	(2)	(3)
Always-New Youth exposure, 1919-22	0.109** (0.049)	0.126** (0.050)	0.117** (0.056)
Not-always-New Youth exposure, 1919-22	0.121* (0.073)	0.125* (0.073)	0.151** (0.071)
New Youth exposure, 1915-1918		-0.078* (0.047)	-0.049 (0.050)
Team FE			Y
Birth year FE			Y
Personal characteristics I			Y
Observations	689	689	689
R-squared	0.388	0.390	0.445

Note. This table separates the exposures to those localities always with the *New Youth* and those with the *New Youth* entry and exit. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1

Table 5. Impacts of Placebo Exposures: Assuming One Were in His Previous or Next Locations
Dependent Variable: CCP=1/0

	Relocation to the next location				Relocation to the previous location			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Teams 1-4				Teams 1-4			
Placebo New Youth Exposure, 1919-22	0.041 (0.046)		0.035 (0.045)	0.041 (0.050)	0.098 (0.068)		0.065 (0.060)	0.018 (0.064)
Placebo New Youth Exposure, 1915-18		0.033 (0.039)	0.024 (0.038)	0.010 (0.049)		0.128* (0.067)	0.108* (0.061)	0.062 (0.063)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y	Y
Team FE	Y	Y	Y	Y	Y	Y	Y	Y
Personal characteristics I	Y	Y	Y	Y	Y	Y	Y	Y
Personal characteristics II				Y				Y
Observations	689	689	689	530	689	689	689	530
R-squared	0.456	0.455	0.457	0.558	0.458	0.460	0.461	0.558

Note. Columns (1)–(4) assume that one were located in the previous prefecture according to his life event records whereas Columns (5)–(8) assume that one were located in the next prefecture.

Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP referees; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit.

Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Comparing with Those Who Just Missed Post-1919
Dependent Variable: CCP=1/0

	(1)	(2)	(3)	(4)
Just who miss post-1919 New Youth	0.024 (0.067)	0.006 (0.067)	0.026 (0.072)	0.016 (0.071)
New Youth exposure, 1919-22	0.123* (0.067)	0.131** (0.066)	0.132* (0.068)	0.136** (0.068)
New Youth exposure, 1915-18		-0.081* (0.049)		-0.053 (0.051)
New Youth exposure, 1919-22 vs. Just-miss	0.100* (0.051)	0.125*** (0.055)	0.105** (0.057)	0.120** (0.059)
Prefecture FE	Y	Y	Y	Y
Team FE			Y	Y
Personal characteristics I			Y	Y
Observations	689	689	689	689
R-squared	0.386	0.389	0.442	0.443

Note. This table compares those with exposure and those who had been to prefectures with post-1919 *New Youth* but missed it due to life events. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP referees. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Other Pro-Communism Periodicals and Activist Groups
Dependent Variable: CCP=1/0

(a) New Youth exposure vs. Two other Communist periodicals

	(1)	(2)	(3)
New Youth, 1919-22 + no other two Communist mag.	0.100*	0.116*	0.142**
	(0.059)	(0.061)	(0.063)
New Youth, 1919-22 + other two Communist mag.	0.113**	0.134**	0.120**
	(0.052)	(0.055)	(0.057)
Prefecture FE	Y	Y	Y
New Youth Exposure, 1915-18		Y	Y
Team FE			Y
Personal characteristics I			Y
Observations	689	689	689
R-squared	0.386	0.389	0.443

(b) New Youth exposure vs. Activist groups

	(1)	(2)	(3)
New Youth, 1919-22 + no Activist org.	0.086*	0.103**	0.110**
	(0.047)	(0.049)	(0.051)
New Youth, 1919-22 + Activist org.	0.169**	0.204***	0.178**
	(0.075)	(0.076)	(0.082)
New Youth Exposure, 1915-18	Y	Y	Y
Prefecture FE		Y	Y
Team FE			Y
Personal characteristics I			Y
Observations	689	689	689
R-squared	0.388	0.392	0.444

Note. This table compares exposure to post-1919 *New Youth* and those also exposed to other Communist magazines and activist groups. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP referees. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

Table 8. Impact on Future Political Outcomes

	(1) 1924-1949		(2) 1924-1949		(3) 1927-1936	
	(1.1) Quit	(1.2) Sacrifice	(2.1) Quit	(2.2) Sacrifice	(3.1) Quit	(3.2) Sacrifice
(1) KMT w. New Youth exposure, 1919-22	0.016 (0.045)	0.152*** (0.052)	0.029 (0.046)	0.147*** (0.053)	-0.006 (0.025)	0.055 (0.038)
(2) CCP w/o New Youth exposure, 1919-22	0.355*** (0.092)	0.303*** (0.106)	0.353*** (0.092)	0.304*** (0.106)	0.580*** (0.053)	0.249*** (0.081)
(3) CCP w. New Youth exposure, 1919-22	0.109* (0.056)	0.554*** (0.065)	0.120** (0.057)	0.550*** (0.065)	0.327*** (0.033)	0.413*** (0.049)
(3)-(2)	-0.245*** (0.100)	0.251** (0.115)	-0.233*** (0.100)	0.246** (0.115)	-0.253*** (0.057)	0.164* (0.087)
New Youth exposure, 1915-18			Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y
Team FE	Y	Y	Y	Y	Y	Y
Personal characteristics I	Y	Y	Y	Y	Y	Y
Observations	689	689	689	689	594	594
R-squared	0.354	0.375	0.356	0.376	0.588	0.417

Note. This table compares the future outcomes of our studied individuals. The reference group is the KMT without post-1919 New Youth exposure.

Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP referees. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix

Table of Contents

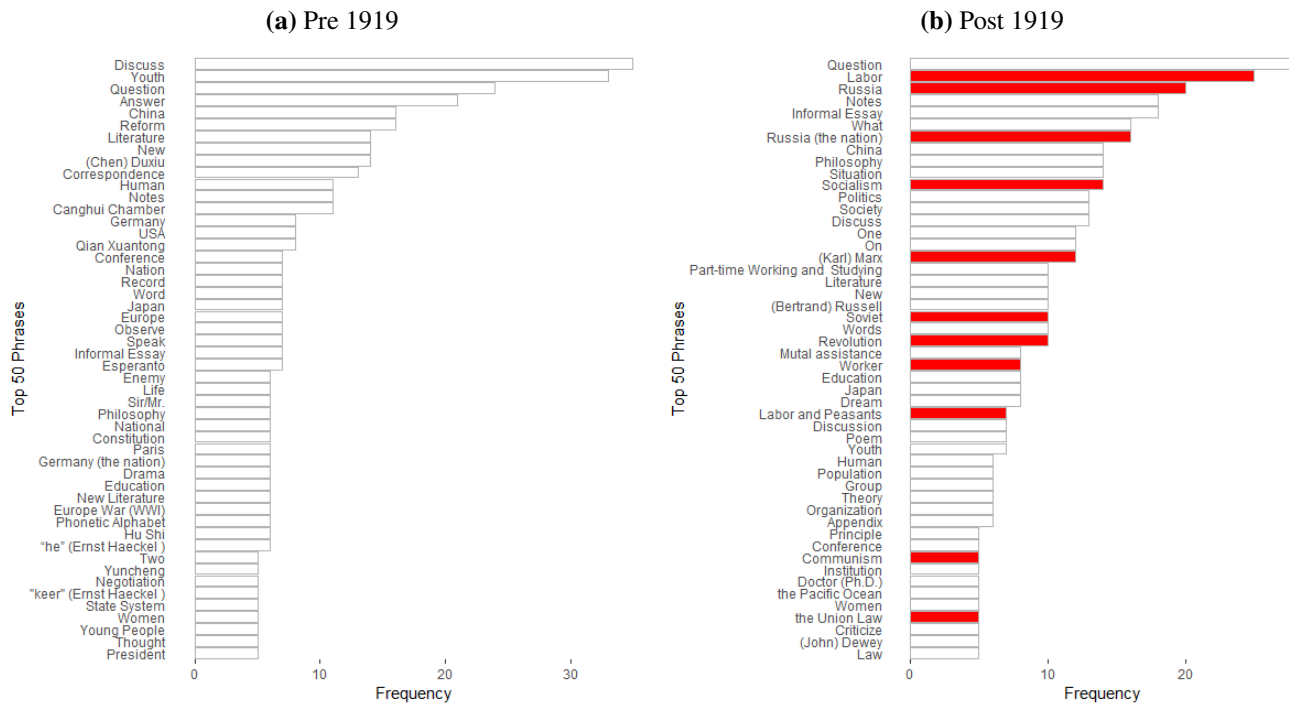
A	Background and Data: More Results	A-2
A.1	Ideological Change of the <i>New Youth</i> : Top words	A-2
A.2	Summary Statistics and Data Sources	A-3
A.3	The Whampoa Cadets: A Biographical Dataset	A-4
A.4	Coding Life History and Exposure to the <i>New Youth</i>	A-8
A.5	The <i>New Youth</i> and Number of Whampoa Cadets	A-13
B	Additional Analyses	A-13
B.1	Heterogeneous Effects across Personal Characteristics	A-13
B.2	Alternative Standard Errors	A-15
B.3	A Continuous Measure of <i>New Youth</i> Exposure	A-16
B.4	Measurement Error	A-16
B.5	Heterogeneous Circulation Locations	A-19

A Background and Data: More Results

A.1 Ideological Change of the *New Youth*: Top words

Figure A.1 plots the top words from the titles of *New Youth* articles. Panel (a) focuses on the pre-1919 articles and panel (b) on post-1919 articles. As shown in Panel (b), after 1919, quite a few words related to Communism stand out, such as labor, Russia, and socialism.

Figure A.1. Top words in the titles pre-post 1919



A.2 Summary Statistics and Data Sources

Table A.2. Data Sources and Summary Statistics

Variables	Data source	Observations	Mean	SD	Min	Max
CCP membership	A	689	0.171	0.377	0	1
Quit	A	689	0.160	0.367	0	1
Sacrifice	A	689	0.241	0.428	0	1
NY exposure, post-19	A, B	689	0.737	0.440	0	1
NY exposure, pre-19	A, B	689	0.552	0.498	0	1
NY exposure, post-19 (County)	A, B	689	0.530	0.499	0	1
NY exposure, pre-19 (County)	A, B	689	0.293	0.456	0	1
Just miss	A, B	689	0.125	0.331	0	1
Other Communist magazines	C	689	0.466	0.499	0	1
Activist Groups	D	689	0.376	0.485	0	1
Team	A	689	3.274	1.736	1	6
Birth year	A	689	1900.7	2.888	1886	1909
Education: middle school and above	A	689	0.956	0.204	0	1
Education: vocational	A	689	0.605	0.489	0	1
Education: college	A	689	0.107	0.310	0	1
#CCP references	A	689	0.180	0.549	0	5
Teams 1-4 only						
Family: gentry	A, E	530	0.249	0.433	0	1
Family: business	A, E	530	0.308	0.462	0	1
Family economic status (1-3)	A, E	530	1.634	0.538	1	3
# Siblings	A, E	530	3.383	2.430	0	16
Married	A, E	530	0.543	0.499	0	1
Having children	A, E	530	0.138	0.345	0	1
Religious	A, E	530	0.109	0.312	0	1
Smoking and drinking	A, E	530	0.051	0.220	0	1

Sources.: A. Chen (2017), *Complete Records on the First-Cohort Cadets of the Whampoa Military School*.

B. *The New Youth*. Reprinted in 1970. Tokyo: Kyukoshoin.

C. Xue (ed. 2014), *Red Collections: Progressive Periodicals from 1915–1949*.

D. Zhang et al. (ed. 1979), *Voluntary Societies during the May Fourth Era*.

E. Central Executive Committee of the Kuomintang (1924), *Detailed Survey for the First to Fourth Teams of the First Cohort of the Military Academy*.

A.3 The Whampoa Cadets: A Biographical Dataset

We use two related sources of data to construct a biographical dataset on the first cohort of Whampoa cadets. First, [Chen \(2017\)](#) compiled the biographies for all 689 cadets. Second, the Whampoa conducted a detailed survey for the 530 cadets that were directly recruited (Teams 1-4). Figure A.3(I) presents an example of the admission questionnaires.

Figure A.3. I. The Admission Questionnaire of Xu Xiangqian

陸軍軍官學校詳細調查表		第 號
姓 名	徐 象 謙	職 別 學 生
年 歲	二十四	
籍貫及住址	山西省五臺縣永安村	
通 信 處	山西省五臺縣東冶鎮寶和店轉	
家 族	父名懋淮兄一人姊二人 母姓趙氏弟一人妹一人子妻姓名	
家庭主要職業	士	
家庭生活狀況	中常	
專門技能或特長	無	
有無烟酒嗜好	無	
有無宗教信仰	孔教	

受過教育	國民師範學校畢業
經過履歷	本縣第一高小學校畢業 山西省立國民師範學校畢業 曾任陽曲縣立第四國民學校教員及川至中學附設小學教員
何時入黨	十三年三月二十九日
人黨介紹人及職業住址	王用賓大本營參議 陳振麟參議員 現住上海霞飛路霞飛巷
何以要入本校	為求軍事知識作將來為本黨工作改進國家之準備
入校介紹人及職業住址	趙連登苗培成 本黨山西代表委員
備 攷	

明說 (一)本表每人填寫一張 (二)除成績備考三欄外均須親筆填寫 (三)本表欄內空處須填明在已見者均須填寫
人數 (四)家庭狀況及地產之有無數量

Figures A.3(II) and (III) provide an example that illustrates our coding process, explained below.

i. CCP Membership and Other Political Outcomes. Based on the cadet's biography, we define three dummy variables to indicate the cadet's CCP membership, whether the cadet quit his party, and whether the cadet died during the war. The biography of Xu Xiangqian shows that he joined the CCP in 1927, and became one of the ten marshals of the People's Liberation Army in 1955. Thus, for Xu Xiangqian, CCP membership=1, Quitting=0, and Sacrifice=0. We highlight the related coding in red in Figure A.3(II).

ii. Personal Characteristics I. We construct four sets of variables to measure the basic personal characteristics from the cadets' biography as shown in Figure A.3(II). First, we code the cadet's birth year. Second, based on the description of the cadet's education experiences, we construct three dummies to indicate whether the cadet received middle school education or above, vocational education, and college education. Third, we list references from the cadet's application to the Whampoa, check the party status of the references, and then calculate how many CCP references the cadet had. Fourth, we code which team the cadet was assigned to. The related coding for Xu Xiangqian is highlighted in blue in Figure A.3(II).

iii. Personal Characteristics II. Based on the survey conducted by Whampoa, we construct three additional sets of variables to further measure an individual's personal characteristics. First, we code the cadets' family background. Specifically, we use two dummies to measure whether the cadet's family had an education tradition, and whether his family run some business. Another category variable is constructed to measure his family economic status: 1 (poor), 2 (middle), to 3 (rich). Second, we collect more demographic information, including the number of siblings, whether the cadet married, and whether the cadet had children. Third, we use two dummies to measure the cadet's religion and behavior traits: whether the cadet believed in some religions, and whether the cadet had drinking and smoking habits. The related coding for Xu Xiangqian is highlighted in Figure A.3(III).

Figure A.3. II. An Example of Coding Based on the Biography



徐向前 (1901~1990)

廣州黃埔中國國民黨陸軍軍官學校第一期畢業，原名象謙，^①別字子敬，後改名向前。山西五台人，五台縣第一高等小學、太原省立國民師範學校速成班畢業，父懸淮，乾涸秀才，早年在內蒙古麻林格爾、涼城教書，母趙金鑾，沒讀過書，勤儉持家，家境中等。胞兄受謙（1899~1975），早年在晉軍當過軍需官，中華人民共和國成立後任太原市人民政府文史研究館館員；妹古月，後改名羅，太原女子師範學校畢業後教書，抗日戰爭爆發後到延安加入中共。自填早年信奉孔教。自填登記處：山西五台縣永安村，通信處：五台縣東冶鎮寶和匠轉交。自填入學前履歷：本縣第一高等小學校畢業，山西省立國民師範學校畢業，曾任陽曲縣立第四國民學校教員及川至中學附設小學教員。

1901年11月8日生於五台縣永安村一個農戶家庭，幼時三年私塾啟蒙，1914年考入東冶鎮陀陽高等小學讀書。因家境不濟，兩年後返回本村讀私塾。1917年至1918年在河北省阜平縣書店裡當書徒，1919年3月考入山西國民師範第一期速成班學習，1921年10月畢業，任陽曲縣立第四國民學校教員，每月薪俸二十塊大洋，任教一個學期即被辭退。後任河邊村川至中學附設小學教員，月俸仍為二十塊大洋，任教四年餘，其間與東冶鎮朱香輝結婚，生有一女松枝。

1924年3月經王用賓（時任廣州大本營參議及奉派北方軍事委員，孫中山指派國民黨一大代表，前中國國民黨本部參議兼北方黨務特派員），陳志鵬（時任國民黨山西省臨時黨部籌備委員，山西省參議會參議，前北京政府參議院參議員、上海國會議員）介紹加入中國國民黨，1924年4月由趙建登（國民黨一大山西省代表，前北京大學文科生參與五四運動，太原國民師範學校教員，山西晚報社社長兼總編輯，國民黨山西省臨時黨部籌備委員）、苗培成（國民黨一大山西省代表，原山西平民中學校長，時任國民黨山西省臨時黨部執行委員兼宣傳部長）無黨投考黃埔軍校，繼往上海參加黃埔軍校招生考試，後至廣州復試，1924年6月考入黃埔中國國民黨陸軍軍官學校第一期第一隊學習，在學期間隨隊北上擔任孫中山副官大本營的警衛工作，1924年11月畢業。

分發黃埔軍校第三期入伍生隊第三隊見習，排長，1925年參加第一次東征和對滇桂軍閥楊希閔部、劉震寰部軍事行動，參加中國青年軍人聯合會活動，在黃埔軍校沒加入中共系因不願作“跨黨分子”。1925年夏派赴謝景翼部國民軍工作，任國民二軍（軍長胡景翼、岳維峻）第六混成旅（旅長弓富魁）教導營教官、旅司令部參謀、第二團少校團附。1926年11月到武漢，任中央軍事政治學校武漢分校南湖學兵團指揮員，軍校學員旅隊政治大隊第一隊少校隊長，工兵大隊大隊長。1927年3月在漢口經樂鵬、梅德鈞介紹加入中共，^②後任國民革命軍第二方面軍總指揮部參謀。1927年9月到廣州，參加廣州起義的準備工作，任工人赤衛隊第六聯隊

Name: Xu Xiangqian
Birth year: 1901

Education (middle or above)=1
Education (vocational)=1
Education (college)=0

CCP membership =1
Time: 1927

隊長。起義失敗後隨部撤退花縣，改編為工農革命軍第四師，任第十團黨代表。1928年1月1日紅四師到海陸豐，任第四師參謀長。1928年6月葉劍英後接任第四師師長，領導廣東海陸豐長裝暴動。

1929年1月根據中共廣東省委指示撤離海陸豐，轉赴上海。1929年6月被中共中央軍委派赴鄂豫邊根據地，任紅軍第十一軍第三十一師師長、代師長，中共鄂豫邊特委委員，鄂豫邊軍事委員會主席，參與領導發展鄂豫邊區紅軍及根據地和指揮反圍剿作戰。1930年2月起任中共鄂豫皖特委委員，紅軍第一軍副軍長兼第一師師長，中共第一軍前敵委員會委員。1931年1月任紅軍第四軍參謀長，1931年7月任紅軍第四軍軍長，1931年11月任紅軍第四方面軍總指揮兼紅四軍軍長，當選為中蘇蘇維埃共和國中央革命軍事委員會委員，參與發展和擴大鄂豫皖邊區紅軍和根據地。1932年10月率紅四方面軍西征，任西北革命軍事委員會副主席，領導創建川陝邊根據地，當選為中蘇蘇維埃第二屆中央執行委員，同中央紅軍會師後，任紅軍前敵總指揮部總指揮。1936年7月紅二、四方面軍會師後，任中共中央西北局委員。1936年10月任中央革命軍事委員會主席團委員，1936年11月任西路軍軍政委員會副主席。

抗日戰爭爆發後，任中共中央軍委委員，國民革命軍第八路軍第一二九師副師長，改裝番號後仍任第十八集團軍第一二九師副師長。^③1939年6月任八路軍山東第一縱隊司令員。1940年12月返回延安，任陝甘寧省綏德軍區副司令員兼參謀長。1941年10月參與籌備組織延安黃埔同學分會，被選進為理事，排名第一。1943年後任抗日軍政大學總校代理校長，1945年5月當選為中共第七屆中央委員。抗日戰爭勝利後，1946年任晉冀魯豫軍區副司令員，華北軍區副司令員兼第一兵團（後改為第十八集團）司令員兼政委。1948年秋兼任華北人民政府委員，後任太原前線司令部司令員兼政委，指揮解放太原戰役。中華人民共和國成立後，任中國人民解放軍總參謀長，人民革命軍事委員會副主席，中共中央軍委副主席，國防委員會第一至三屆副主席，第二、三屆全國人大常委會副委員長，國務院副總理兼國防部部長，中共第八屆中央政治局委員，第九、十屆中央委員，第十一、十二屆中央政治局委員。

1955年9月27日被授予中華人民共和國元帥軍銜。1984年6月任黃埔軍校同學會會長，1988年2月任黃埔軍校同學會名譽會長，1990年5月10日在北京馬所接見從臺灣回大陸觀光的一期同學鄧文儀等，實現了海峽兩岸黃埔生和解。1988年主動辭去中央軍委副主席、中華人民共和國國務院副總理等職。1990年9月21日因病在北京逝世。著有《歷史的回顧》、《徐向前軍事文集》等。1988年被中華人民共和國中央軍事委員會確定為中國人民解放軍軍事家，是中國人民解放軍創建人和領導者之一。^④其子徐小岩，歷任中國人民解放軍總參謀部通信部部長，南京軍區副司令員，中國人民解放軍總裝備部科技委員會副主任，中將軍銜。2010年聘任中央新聞影片廠拍攝的大型歷史文獻紀錄片《黃埔軍校》攝製組歷史顧問。

References: Zhao Liandong, Miao Peichen

Team no.=1

315

316

Political outcomes:
Quit=0
Sacrifice=0

A-6

①、自廣州黃埔軍校第一期畢業生詳見《黃埔軍校第一期第一隊學員名錄》。

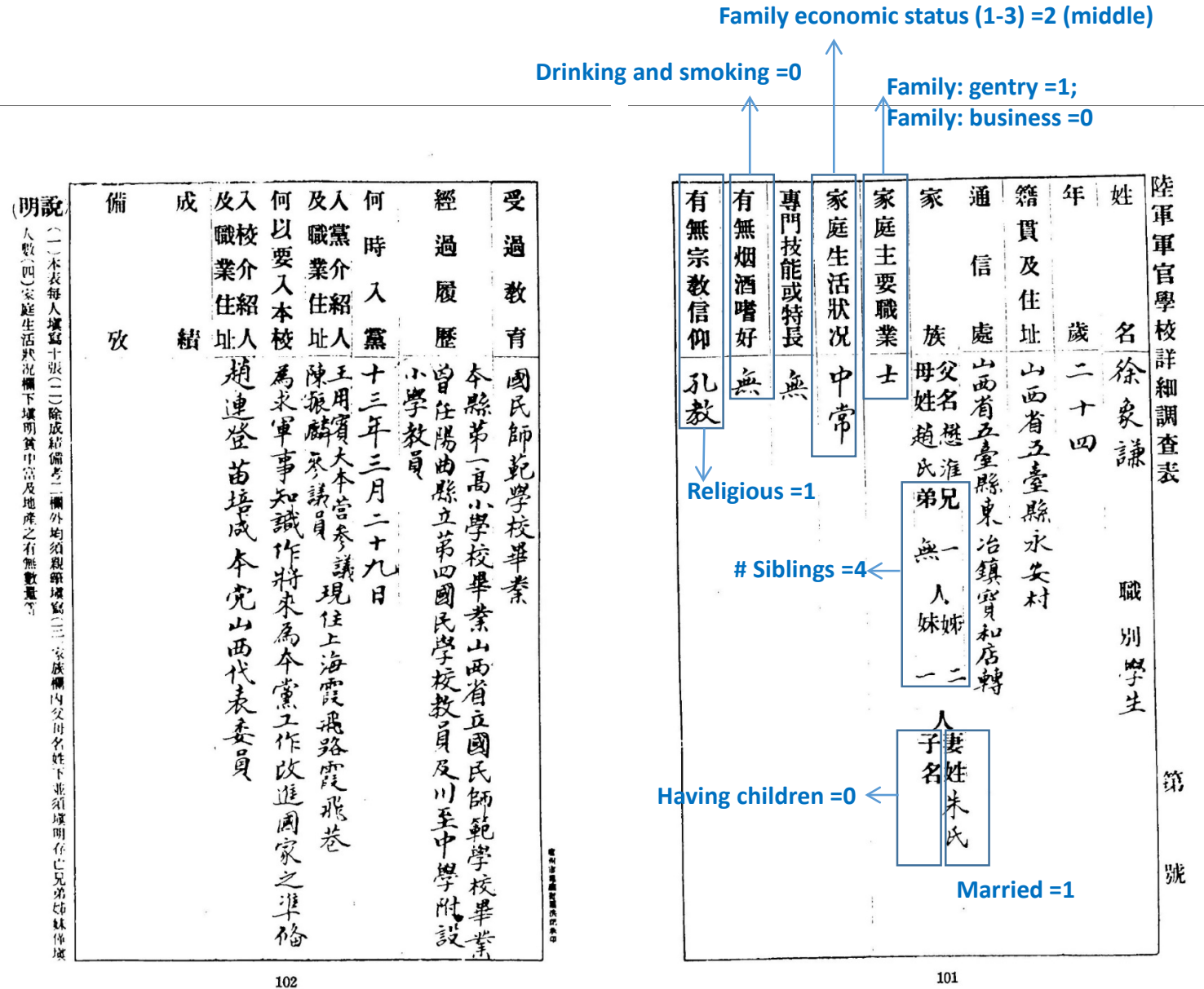
②、謝景翼部國民軍工作，任國民二軍（軍長胡景翼、岳維峻）第六混成旅（旅長弓富魁）教導營教官、旅司令部參謀、第二團少校團附。1926年11月到武漢，任中央軍事政治學校武漢分校南湖學兵團指揮員，軍校學員旅隊政治大隊第一隊少校隊長，工兵大隊大隊長。1927年3月在漢口經樂鵬、梅德鈞介紹加入中共，^②後任國民革命軍第二方面軍總指揮部參謀。1927年9月到廣州，參加廣州起義的準備工作，任工人赤衛隊第六聯隊

③、徐向前，解放軍出版社1988年4月《歷史的回顧》第29頁記載。

④、鄧知聖主編，軍史雜誌社1999年10月15日《中國人物小傳》第十八輯記載。

⑤、廖雲生主編，中共中央文獻出版社2001年6月《中國共產黨歷史大辭典》增訂本第419頁記載。

Figure A.3. III. An Example of Coding Based on the Detailed Survey



A.4 Coding Life History and Exposure to the *New Youth*

We take three steps to define and code whether an individual cadet was exposed to the *New Youth* and Communist ideology, as explained below.

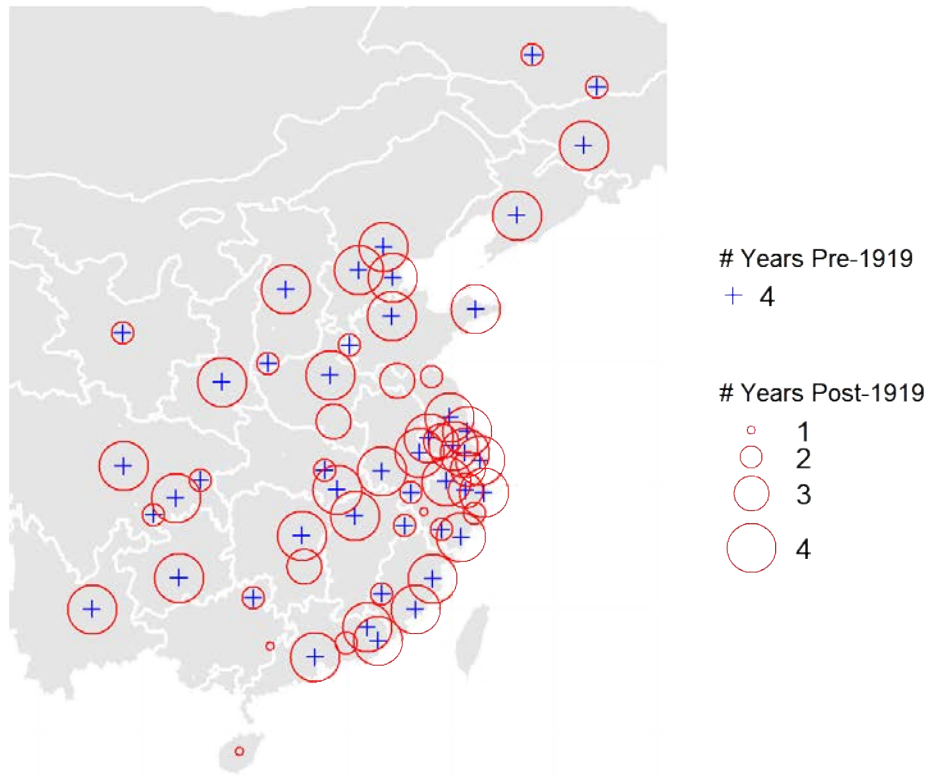
Step 1. Coding the Circulation of the *New Youth* Over Time The *New Youth* magazine listed its circulation locations in each volume. Figure A.4(I) provides an example.

Figure A.4. I. An Example of the Circulation of the *New Youth*

各埠代派處																									
重慶	重慶	成都	成都	成都	蘭州	貴州	雲南	西安	西安	西安	連雲港	太原	煙台	濟南	濟南	大名	保定	保定	天津	北京	北京				
崇文書局	二酉山莊	觀石齋	點石齋	崇文書局	正山書房	翠明社	維新書局	正誼公司	新智識圖書社	公益書局	交通義書社	晉新書局	教育圖書局	山東官書局	大名官書局	翠玉山房	直隸書局	新華書局	直隸書局	浚花書局	龍文書局				
蘇州	蘇州	坎市	廣門	福州	福州	屯溪	蕪湖	安慶	廣信	南昌	南昌	武昌	漢口	漢口	長沙	嘉應州	桂林	汕頭	廣州	廣州	開封	開封	梁山	德州	
瑪瑙房	振新書社	廣益書局	新民書社	陳文書莊	宏開書莊	科學圖書社	科學圖書社	萬智書局	益智書局	點石齋	開明書局	昌明書局	會文書局	昌明書局	華益圖書公司	啓新書局	石渠書局	共和編譯局	蒙學編譯局	共和編譯局	百城書館	文會山房	益山房	二酉山房	
新加坡	新加坡	哈爾濱	龍江	吉林	奉天	奉天	奉天	奉天	溫州	溫州	處州	紹興	甯波	甯波	杭州	常州	揚州	南通州	南通州	無錫	無錫	南京	南京	蘇州	蘇州
普益印務公司	曹萬豐書莊	魁新書堂	維新書局	永衡官書局	關東書局	神州圖書行	會文書局	圖書發行所	維新書局	日新書局	悅興書局	教育書局	文興書局	問明書局	晉經書局	競存書局	聚生書局	導文書局	樂群書局	文華書局	新民圖書館	共和書局	小怡福林	文怡福林	

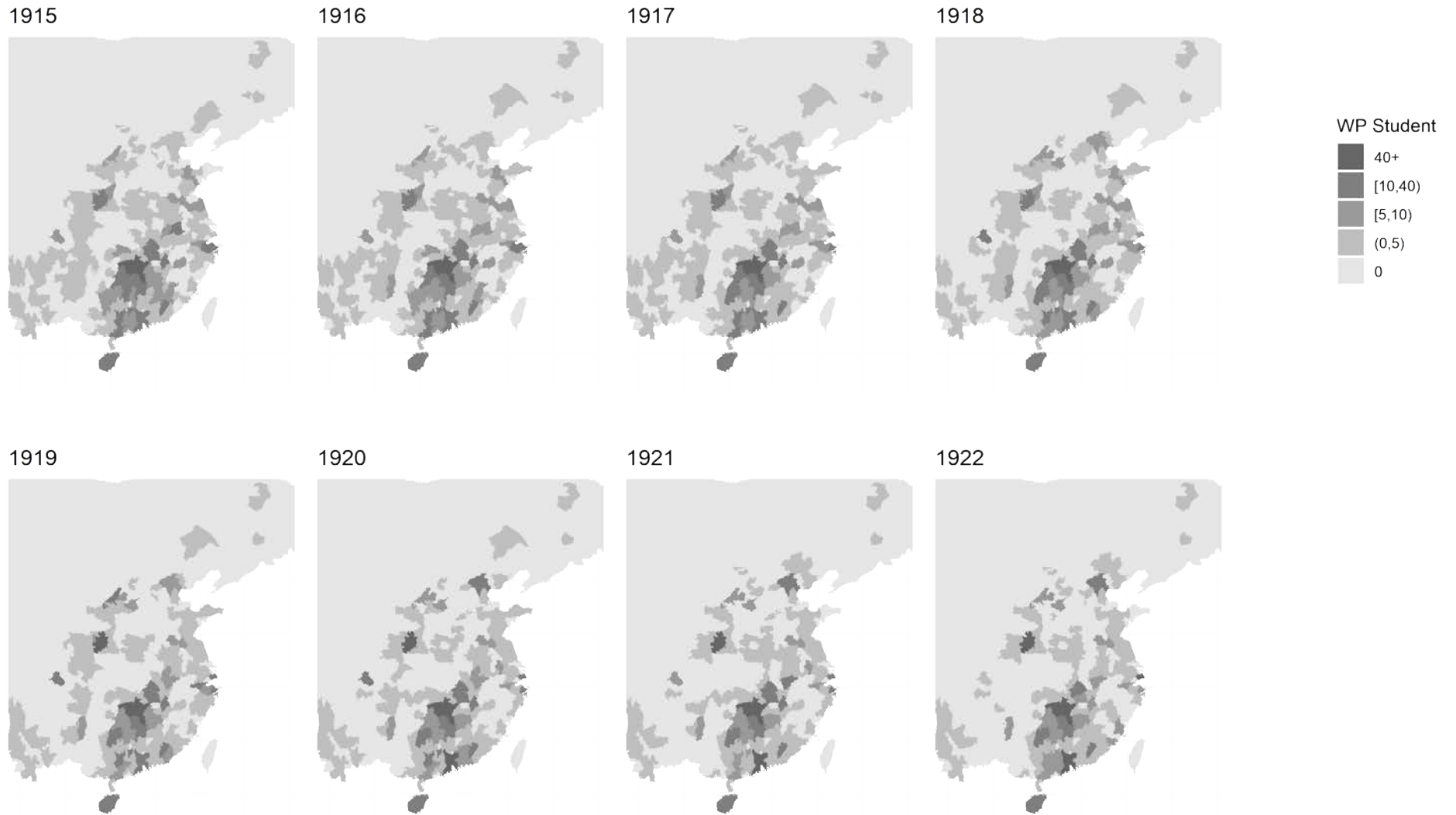
We code the circulation locations for all volumes during 1915–22. Figure A.4(II) presents the geographical circulation of the *New Youth* over time. As shown, the *New Youth* was always available for a group of prefectures whereas there existed entry and exit in other prefectures. Not surprisingly, the former group tends to be larger cities.

Figure A.4. II. Circulation Locations of the New Youth



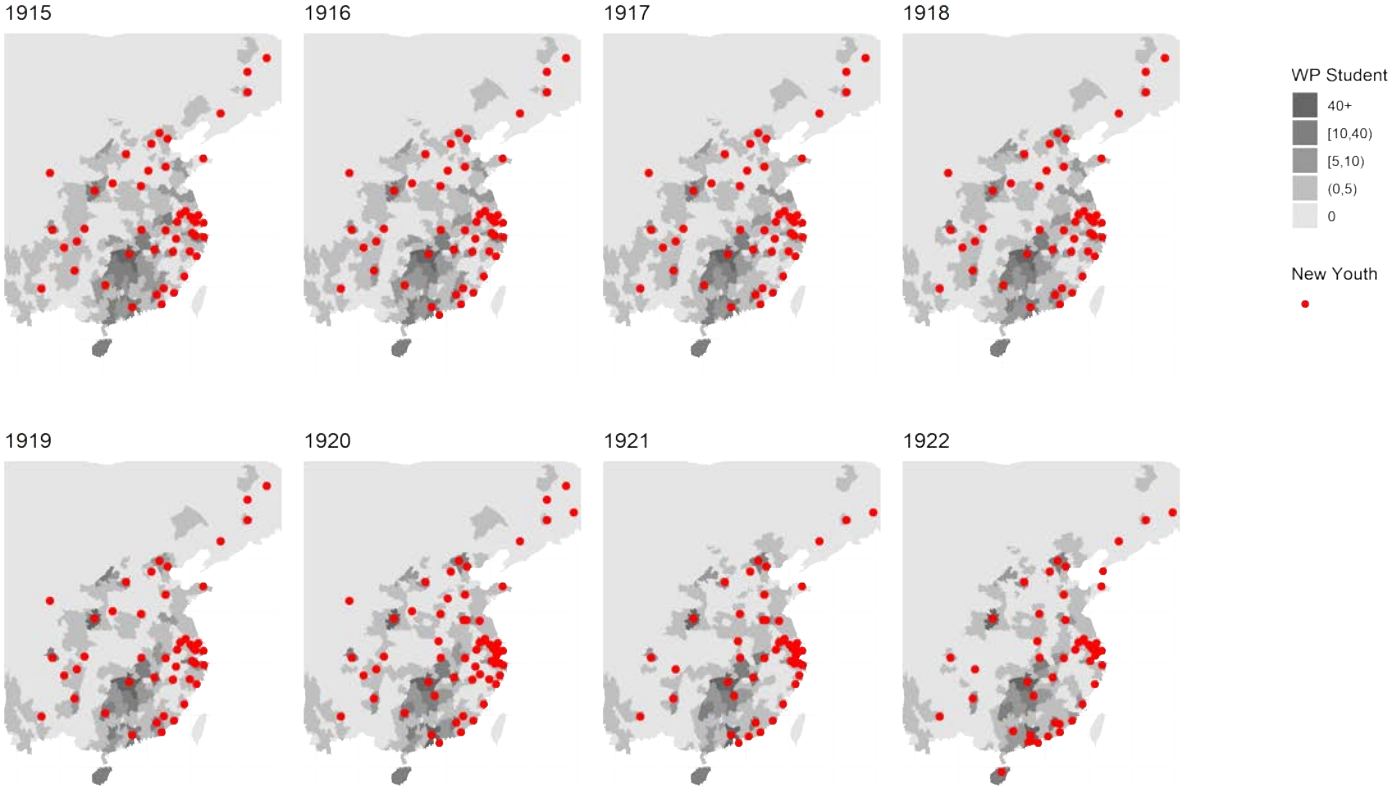
Step 2. Coding the Locations of the Cadets Over Time. Based on the cadets’ biography, we use the information on their recorded life events, focusing on time and location to code the cadets’ location in each year, as illustrated in Figure A.4(III). This way, we construct an individual-year-level panel dataset with the cadets’ location. Figure A.4(IV) presents the geographical distribution of all cadets in each year during 1915-22. As shown, they were widely spread across China.

Figure A.4. IV. Cadet Location Year by Year



Step 3. Defining Exposure to the *New Youth* Finally, we define whether a cadet was exposed to the *New Youth* and the Communist ideology. We combine cadets' location and *New Youth*'s circulation in Figure A.4(V). If a cadet lived in a locality (represented by the gray background) with the *New Youth* circulation (represented by the red point), this cadet was exposed to the *New Youth*. Using Xu Xiangqian as an example, Xu lived in Taiyuan during 1919–22. Because the *New Youth* was available in this location and during this time, Xu was exposed to the *New Youth*.

Figure A.4. V. Exposure to the *New Youth*



A.5 The *New Youth* and Number of Whampoa Cadets

Our analyses focus on examining how the *New Youth* exposure affects the party choices of the Whampoa cadets. We also wonder whether the *New Youth* exposure affects how many Whampoa cadets enter our sample. To check this relationship, we build a prefecture-level panel for two periods (1915–1918 and 1919–1922). Each observation in the panel data indicates the number of Whampoa cadets who had been to a prefecture in that period. Without any controls, there is a positive association between the *New Youth*'s exposure and the number of Whampoa cadets (Column (1) of Table A.5). However, this relationship seems to be driven by the size of a locality: Once we control for city size (i.e., the size of urban population in 1920), the positive association disappears (Column (2)). In Columns (3)–(4), we control prefecture fixed effects and find that the change in *New Youth* exposure is not significantly correlated with a change in the number of Whampoa cadets either.

Table A.5. The Impacts of Ideology Exposure across Localities
Dependent Var.: Number of Whampoa Cadets

	(1)	(2)	(3)	(4)
<i>New Youth</i> circulation	4.334*** (1.397)	0.863 (0.806)	0.119 (0.710)	-0.316 (0.774)
City population		0.011*** (0.003)		
City population*Post1919		0.011*** (0.002)		0.011*** (0.002)
Post1919	0.018 (0.183)	0.093 (0.127)	0.154 (0.194)	0.130 (0.126)
Prefecture FE			Y	Y
Observations	686	686	686	686
R-squared	0.141	0.294	0.0357	0.549

Note. Standard errors in parentheses are clustered at the prefecture level; *** p<0.01, ** p<0.05, * p<0.1.

B Additional Analyses

B.1 Heterogeneous Effects across Personal Characteristics

Table B.1 presents the interaction effects between post-1919 *New Youth* exposure and various personal characteristics. While there seems to be some complementarity between social networks and ideology exposure, there are no strong heterogeneous patterns in general.

Table B.1. Impact of Post-1919 *New Youth*: Heterogeneous Effects. Dependent Variable: CCP=0/1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NY exposure, 1919-22	0.122** (0.055)	0.112* (0.058)	0.118** (0.055)	0.126** (0.055)	0.117** (0.056)	0.125** (0.056)	0.117** (0.058)	0.117** (0.056)	0.116** (0.056)	0.117** (0.056)	0.118** (0.055)	0.118** (0.056)
NY*1(Birth year>1900)	-0.116 (0.095)											
NY*Education: vocational		-0.075 (0.110)										
NY*Education: college			0.019 (0.132)									
NY*# CCP referees				0.135* (0.076)								
NY*Family: gentry					0.049 (0.087)							
NY*Family: business						0.151* (0.091)						
NY*Family economic status							0.003 (0.085)					
NY*# siblings								0.000 (0.014)				
NY*Married									0.010 (0.090)			
NY*Having child										-0.022 (0.129)		
NY*Religious											0.044 (0.116)	
NY*Smoking and drinking												0.071 (0.146)
NY exposure, 1915-18	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Prefecture FE, Team FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Personal charar. I, II	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	530	530	530	530	530	530	530	530	530	530	530	530
R-squared	0.545	0.543	0.542	0.546	0.543	0.546	0.542	0.542	0.542	0.542	0.542	0.542

Note. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

B.2 Alternative Standard Errors

Robust standard errors are reported in parentheses, whereas two-way clustering standard errors at home county level and birth year level are presented in brackets. These standard errors are close to those clustered at the home county level.

Table B.2. Impact of Post-1919 *New Youth*: Various Standard Errors
Dependent Variable: CCP=0/1

New Youth Exposure defined at	Prefecture level						County level					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Teams 1-4						Teams 1-4	
New Youth Exposure, 1919-22	0.109 (0.042)*** [0.046]**		0.128 (0.044)*** [0.048]**	0.127 (0.047)*** [0.050]**	0.127 (0.054)** [0.053]**	0.117 (0.053)** [0.051]**	0.153 (0.064)** [0.057]**		0.152 (0.064)** [0.057]**	0.181 (0.070)*** [0.053]***	0.162 (0.081)** [0.048]***	0.172 (0.083)** [0.046]***
New Youth Exposure, 1915-18		-0.047 (0.046) [0.042]	-0.081 (0.048)* [0.040]*	-0.055 (0.051) [0.044]	-0.091 (0.059) [0.052]	-0.094 (0.061) [0.050]*		-0.044 (0.080) [0.106]	-0.038 (0.079) [0.107]	-0.006 (0.085) [0.102]	-0.103 (0.101) [0.110]	-0.104 (0.103) [0.109]
Prefecture FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
County FE							Y	Y	Y	Y	Y	Y
Team FE				Y	Y	Y				Y	Y	Y
Personal characteristics I				Y	Y	Y				Y	Y	Y
Personal characteristics II						Y						Y
Observations	689	689	689	689	530	530	689	689	689	689	530	530
R-squared	0.386	0.381	0.389	0.443	0.534	0.542	0.560	0.552	0.560	0.603	0.747	0.750

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP referees; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit. Robust standard errors are presented in parentheses and two-way clustering standard errors at the hometown level and the birth year level in brackets. *** p<0.01, ** p<0.05, * p<0.1.

B.3 A Continuous Measure of *New Youth* Exposure

Our finding is robust to using a continuous measure of *New Youth* exposure, where we calculate the exposure as the share of years being exposed. The results are presented in Table B.3.

Table B.3. A Continuous Measure of New Youth Exposure
Dependent Variable: CCP=0/1

	(1)	(2)
	Teams 1-4	
Average New Youth exposure, 1919-22	0.159*** (0.058)	0.161** (0.070)
Average New Youth exposure, 1915-18	-0.002 (0.069)	-0.068 (0.081)
Prefecture FE	Y	Y
Team FE	Y	Y
Personal characteristics I	Y	Y
Personal characteristics II		Y
Observations	689	530
R-squared	0.464	0.564

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP references; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

B.4 Measurement Error

We conduct three sets of analyses to make sure our findings are not driven by measurement error at the locality or individual levels.

i. Dropping Metropolitan Cities. Our estimates are similar or a bit larger than our baseline estimate if we drop the metropolitan cities from our sample.

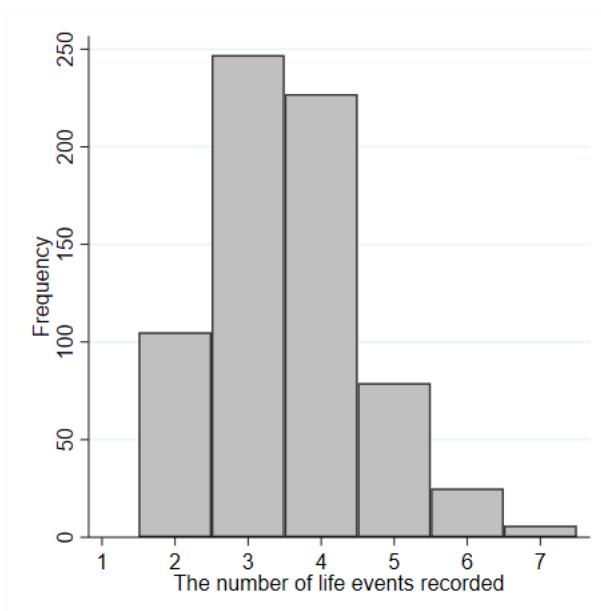
Table B.4. I. Dropping Big Cities
Dependent Variable: CCP=0/1

	All cadets				Teams 1-4			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Excluding who had been CCP mean:	Beijing 0.163	Shanghai 0.163	Guangzhou 0.177	All three 0.172	Beijing 0.160	Shanghai 0.160	Guangzhou 0.179	All three 0.173
New Youth Exposure, 1919-22	0.135*** (0.047)	0.133*** (0.045)	0.132** (0.066)	0.212** (0.088)	0.127** (0.058)	0.126** (0.054)	0.140** (0.069)	0.199** (0.088)
New Youth Exposure, 1915-18	Y	Y	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y	Y	Y
Team FE	Y	Y	Y	Y	Y	Y	Y	Y
Personal char. I	Y	Y	Y	Y	Y	Y	Y	Y
Personal char. II					Y	Y	Y	Y
Observations	613	652	417	314	456	494	407	306
R-squared	0.438	0.445	0.592	0.563	0.547	0.558	0.602	0.577

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP references; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

ii. Precision of Life Event Records. The major life events for our studied individuals—in their early 20s in 1924—are education and jobs. On average, each cadet has 3.6 recorded life event records. As Figure B.4 shows, the majority of the cadets (474 out of 689) have 3 or 4 recorded events. About 15% (105 out of 689) have 2 recorded events, and 16% (110 out of 689) have at least 5 recorded events.

Figure B.4. The Number of Life Events Recorded



To check whether our finding is affected by the number of events recorded, we examine whether our estimate varies by the number of records. This is not the case, as shown in Table B.4(II). Our finding also holds if we remove those who have few or many events recorded.

Table B.4. II. Considering the Number of Recorded Life Events
Dependent Var.: CCP=0/1

	(1)	(2)	(3)	(4)	(5)	(6)
					Teams 1-4	
#Recorded events:	>2		<5		>2	
					<5	
New Youth exposure, 1919-22	0.127**	0.165***	0.166***	0.112**	0.135**	0.153**
	(0.053)	(0.060)	(0.058)	(0.055)	(0.061)	(0.063)
#Recorded events * New Youth, 1919-22	0.011			-0.002		
	(0.045)			(0.049)		
#Recorded events:	0.052			0.065		
	(0.036)			(0.045)		
New Youth exposure, 1915-18	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y
Team FE	Y	Y	Y	Y	Y	Y
Personal characteristics I	Y	Y	Y	Y	Y	Y
Personal characteristics II				Y	Y	Y
Observations	689	584	579	530	488	426
R-squared	0.455	0.489	0.448	0.554	0.574	0.599

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP references; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

iii. Missing Years of Recorded Events. Some recorded events did not include exact years. In these cases, we use the beginning year of the next event to infer the ending year or infer the years based on the school system. These assumptions may induce some measurement error. To check the validity of our method, we apply the same inference method to all our sample by pretending that we only knew the birth year, end year, and life events for the studied individuals.

Columns (1) and (3) of Table B.4 present our original estimates and Columns (2) and (4) with the inferred years for everyone. The similar estimates confirm that our findings are not driven by our inference method.

Table B.4. III. Validating the Inference Method
Dependent Var.: CCP=0/1

	(1)	(2)	(3)	(4)
			Teams 1-4	
New Youth exposure, 1919-22	0.127** (0.052)		0.117** (0.056)	
New Youth exposure, 1919-22 (Based on events only)		0.110** (0.056)		0.123** (0.058)
NY exposure, 1915-18	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y
Team FE	Y	Y	Y	Y
Personal characteristics I	Y	Y	Y	Y
Personal characteristics II			Y	Y
Observations	689	689	530	530
R-squared	0.443	0.440	0.542	0.542

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP references; Personal characteristics II include family backgrounds (gentry, business), family economic status (1-3), number of siblings, marriage status, a dummy indicating whether they have children, religious belief, and smoking and drinking habit. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.

B.5 Heterogeneous Circulation Locations

To gauge the reading rate, we examine heterogeneity across circulation locations in three ways. First, we separate the prefecture-level post-1919 *New Youth* into a breakdown of within a county and outside the county. Columns (1)–(2) of Table B.5 show little impact if the post-1919 *New Youth* was available outside of a cadet’s located county.

Second, we separate the circulation county into small and big cities, defined by a threshold of urban population size of 100,000. Using the same specification as our baseline, we obtain estimates of 0.26 for small cities and 0.13 for big cities (Columns (3)–(4)).

Finally, a subsample of the circulation points were bookstores located in schools. Based on the life history of the cadets, we know whether a cadet was studying in those schools. When separating this very local exposure from others, we find that the coefficient of this very local exposure is 0.26 whereas that of the rest is 0.15 (Columns (5)–(6)).

Table B.5. The Impacts of Ideology Exposure across Localities
 Dependent Var.: CCP=0/1

	(1)	(2)	(3)	(4)	(5)	(6)
New Youth (NY) Exposure, 1919-22 (pref. level)						
different county with the NY	0.040 (0.050)	0.051 (0.054)	0.038 (0.049)	0.049 (0.054)	0.032 (0.050)	0.044 (0.054)
same county with the NY	0.160*** (0.056)	0.161*** (0.057)				
same county with the NY: small city			0.251*** (0.088)	0.260*** (0.080)		
same county with the NY: big city			0.135** (0.059)	0.131** (0.061)		
same county with the NY: same school					0.250* (0.149)	0.263* (0.155)
same county with the NY: others					0.151*** (0.058)	0.152*** (0.058)
NY Exposure, 1915-18	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y
Team FE		Y		Y		Y
Birth year FE		Y		Y		Y
Personal characteristics I		Y		Y		Y
Observations	689	689	689	689	689	689
R-squared	0.400	0.451	0.403	0.455	0.400	0.452

Note. Personal characteristics I include birth year fixed effects, vocational education, college, and number of CCP references. Standard errors in parentheses are clustered at the home county level; *** p<0.01, ** p<0.05, * p<0.1.