UC Riverside
UC Riverside Electronic Theses and Dissertations

Title
Crises and Socio-Economic Development

Permalink
https://escholarship.org/uc/item/9wv2k4k0

Author
Gong, Da

Publication Date
2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Crises and Socio-Economic Development

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Da Gong

June 2024

Dissertation Committee:

- Dr. Joseph R. Cummins, Chairperson
- Dr. Sarojini Hirshleifer
- Dr. Ugo Antonio Troiano
- Dr. Yang Xie

Copyright by
Da Gong
2024

The Dissertation of Da Gong is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to express my sincere gratitude to my advisors and dissertation committee members, Dr. Joseph R. Cummins, Dr. Sarojini Hirshleifer, Dr. Ugo Antonio Troiano, and Dr. Yang Xie, for their invaluable support and guidance in my research and career development. Their dedicated help and constructive advising have helped me grow from a student into a PhD. They set a great role model for me through their dedication, expertise, and unwavering support, inspiring me to strive for excellence in my own academic endeavors.

I would also like to express my appreciation to Dr. Kevin M. Esterling, Dr. Michael Bates, Dr. Wei Zhao, and all the other professors and researchers who provided support and comments on my research, served on my oral qualifying exam, and offered invaluable help and suggestions during my job search.

I am also grateful for the tremendous support from Anna, Gary Kuzas, Jason Chou, my colleagues at GradQuant and GradSuccess, as well as my friends, classmates, coauthors, and UCR alumni, for their help in my study, research, and job search.

To my beloved family.

To all these sleepless nights.

ABSTRACT OF THE DISSERTATION

Crises and Socio-Economic Development

by

Da Gong

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, June 2024
Dr. Joseph R. Cummins, Chairperson

This dissertation consists of three independent essays with focus on crises and socio-economic development that sit at the intersection of Political Economy, Development Economics, and Public Economics.

Chapter 1 examines how the impact of traumatic experiences on contemporary trust could vary across different initial social capital levels within the context of the Confucian clan and the Great Chinese Famine.

Chapter 2 offers the first comprehensive examination of the economic impacts of China's zero-COVID policy. To achieve this, we utilize an original panel dataset that includes county-level data on daily COVID risk levels.

Chapter 3 studies the effect of China's anti-contagious policy on labor market outcomes in 2020 by exploiting variation in the duration of the zero-Covid policy in China, which is triggered by the outbreak of new cases of COVID-19 in a 14-day observation window.

Contents

List of Figures

List of Tables

Chapter 1

Formation and Evolution of Beliefs: Famine Experience and Trust in Neighbors

1.1 Introduction

Social capital is associated with economic development, institutions and trade.[1] Understanding the dynamics of social capital accumulation and dissipation can have important economic and political implications, but the literature has yielded mixed findings, with no consensus on the impact of negative shocks on social capital formation.[2] Although (123) and (70) conjecture that different initial conditions may lead to the evolution of social capital into equilibriums of mistrust or trust, they provide limited empirical evidence .

In this paper, we provide evidence to shed light on this puzzle by estimating how the impact of traumatic experiences on one aspect of social capital — contemporary trust — could vary across different initial social capital levels. We hypothesize that when individuals

[1](11; 113; 124; 5; 111; 72; 88; 39; 4; 71)

[2]A strand of literature documents that negative shocks, such as slave trade (105), political repression (134) and wars (44) lead to a decrease in social capital. By contrast, another strand find that exposure to adverse events intensify the strength of social capitals (23; 15; 16).

derive substantial benefits from the initial conditions of social capital within a community in the face of negative shocks, it strengthens trust among the community members.

To examine the hypothesis, we focus on a kinship-based historical institution in China – the Confucian clan, where social capital is deeply embedded, in the context of the Great Chinese Famine ["the Famine" henceforth] (1959-1961). We begin by utilizing trust data from the China Family Panel Survey (CFPS) to examine the impact of the Famine on contemporary trust. To achieve this, we exploit differential exposure to county-level famine intensity across cohorts before and after the Famine. Then, we employ historical data on social capital, specifically pre-famine clan strength, to implement a triple difference design. This allows us to compare the famine effects in counties with high clan strength to those in counties with low clan strength. Through this analysis, we identify the impact of initial conditions on the relationship between negative shock and contemporary trust.

We measure the clan strength of each county by using the density of pre-PRC[3] genealogies—books that record lineages' male members, clan rules and moral obligations for members since its inception. In China's rural society, the Confucian clan, a predominant lineage organization, served as a risk-sharing and resource-pooling institution for thousands of years, making it a fundamental form of social capital (38). Genealogy books serve to promote clan solidarity, facilitate intra-clan intertemporal exchange, and function as a fundamental aspect of clan activities while also serving as an indicator of clan cohesion. (29; 39).

As a risk-sharing institution, the clan played a pivotal role throughout China's history during periods of weather shocks,[4] and notably, it also saved millions of lives during the Famine, known as deadliest famine ever recorded in human history. During Mao's Great Leap Forward era, a series of inflexible and progressive government procurement policies, along with systematic misallocation of food, led to the Great Chinese Famine from 1959 to 1961. It is estimated that as many as 30 million people died during this period,

[3] the People's Republic of China

[4] (37) documents that a 10% increase in clan density reduces the freqency of severe drought induced cannibalism by 4.78%

with 85 percent of Chinese counties affected. Despite the non-negotiable orders from the central government, as noted by (29), local clans empowered peasants to organize and resist excessive grain procurement from higher-level authorities, and facilitate intra-clan lending. According to their findings, one standard deviation increase in the clan strength is associated with a reduction of 1.45 to 1.61 deaths per thousand people during the Famine years.

We measure the famine intensity at the county level using the excess mortality rate during the years of the Famine. The mortality data is manually collected from compilation of statistics, local government reports, and county gazetteers.[5] One concern is that the excess mortality rate may be correlated with clan strength, potentially biasing our estimation. To address this concern, we employ weather shocks during the Famine and soil suitability for grain production as instrumental variables (IV) for famine intensity. These IVs are unlikely to correlated with initial clan strength. While long-term climate variability has been found to causally determine social capital (23; 61), the weather shock during the specific years 1959 to 1961 is not expected to be correlated with local clan strength. Furthermore, although a strand of literature finds that wetland rice farming has persistence influence on social norms and cooperative behavior (126; 125; 138), the procurement policy does not show a preference for rice over wheat or other grains and is therefore unlikely to be correlated with pre-famine clan culture.

We find that the Famine results in an average increase of 0.75 points in trust scores among the subgroup with high initial clan strength, which is a noteworthy point estimate, considering that the average trust score among neighbors is 6.5. In the context of the triple-differences specification, we find that the traumatic experience strengthens trust among those who received support from their clans, leading to an increase in trust scores ranging from 0.42 to 0.66 points. Our findings remain stable to individual level controls, different levels of fixed effects, alternative famine measurements and alternative clan measurements. Our results also remain robust to instrumental variable estimation. The event study estimates validate the parallel trends assumption.

[5] Please click this link to see details of China Gazetteer Project.

To investigate whether the development of trust in neighbors is influenced by the framework proposed in our study, we perform a placebo test and replicate our main analysis by assessing the impacts on generalized trust and trust in parents. Our findings indicate that the impact of famine on generalized trust and trust in parents does not vary across initial clan conditions. This supports our hypothesis that clan culture likely played a pivotal role in safeguarding individuals during the famine and therefore reshaping their beliefs.

Our work contributes to several strands of literature. Firstly, it contributes to the growing body of research on trust formation and cultural persistence. Existing studies present mixed findings, demonstrating that historical negative shocks can either undermine or foster social capital and related behaviors. For instance, social capital may be undermined by slave trade (105), political repression (134), or wars (44). On the other side, cooperation and civil participation could also be induced by weather shocks (23), earthquakes (15) or civil wars (16). We contribute to this literature by understanding the role of initial conditions of social capital in steering trust towards different self-enforcing equilibrium. Additionally, our empirical findings complement the theoretical frameworks proposed by (123) and (70).

Secondly, our research contributes to the existing literature on cultural and institutional bifurcation (68; 69; 4). Unique initial cultural conditions lead to the emergence of diverse social organizations, while their subsequent proliferation strengthens their distinct cultural traits (68). Our findings support this view by illustrating a self-enforcing cultural traits—kinship-based interactions reinforce trust towards kin under negative shocks. Specifically, we extend this argument by delving deeper into the impact of a particular historical shock—the Great Chinese Famine. This exploration aims to establish a clear causal link between culture and institution. More broadly, our paper provides insights into the reasons behind certain nations being trapped in a vicious cycle of extractive institutions, resulting in less development (2). Social capital rooted in "limited morality" may foster cooperation within family members but simultaneously hinder the development of inclusive institutions, thereby impeding economic growth (4; 39). The self-reinforcing characteristic of "limited morality" societies, such as kinship-based networks, elucidates the challenge these nations face in developing modern institutions without external shocks (1).[6]

Thirdly, our work also relates to the literature on the the role of kinship-based networks as risk-sharing institutions, particularly in the contexts of China and Sub-Saharan Africa (135; 48; 137; 51; 102; 103; 104). (29) reveals that clans play a crucial role in mitigating famine intensity through informal lending and collective resistance. We extend this argument by revealing that risk-sharing function within kinship-based network is strengthened through survival experiences.

The rest of the paper is organized as follows. Section 2 provides conceptual framework and background to guide the empirical analysis. Section 3 describe the data. Section 4 presents the empirical strategy and results. Section 5 presents the instrumental variables estimation. Section 6 provides Robustness Checks. Section 7 concludes.

1.2 Conceptual Framework and Background

In this section, we provide a brief conceptual framework and a historical background to clan culture and the Great Chinese Famine.

1.2.1 Conceptual Framework

Social capital refers to the attitudes, beliefs, norms and values that support cooperation (70). Mixed findings in the literature suggest that historical negative shocks have the potential to either destroy or cultivate trust and related behaviors. (105) finds that individuals from ethnic groups with significant exposure to the slave trade tend to demonstrate lower levels of trust in their relatives and neighbors. This phenomenon can be

[6] Literature on relationship between culture and institutions document that "generalized morality" is an important factor that enforce cooperation between unrelated individuals, thereby boosting economic growth (113; 124; 111). However, if social capital is rooted in "limited morality", it may foster cooperation within clan members while impeding broader societal development (4; 39). The self-enforce cultural traits in "limited morality" could contribute to explaining the "the Great Divergence" and "Narrow Corridor" pattern (110; 2). Kinship-based networks, like clans, may mitigate shocks and contribute to the prosperity of agrarian-based economies. However, they can also pose obstacles to the development of modern institutions in these countries.

attributed to the historical practice of individuals being frequently sold into slavery by people in their own communities, including neighbors and even family members. Other negative shocks such as political repression (134), or wars (44) have been found to have persistent negative impacts on cooperation and trust. In contrast, (23) finds that regions experienced more frequent climate-related risks exhibit enhanced cooperation among neighboring communities, facilitated by mutual insurance. Similarly, negative shocks such as earthquakes (15) and civil wars (16) have been identified as reinforcing local social capital.

To reconcile the mixed findings in the literature, we propose a brief conceptual framework guided by (123) and (70). We propose that in the presence of a negative shock, if the net benefits of cooperation are sufficiently high, the society will naturally self-enforce towards an equilibrium of trust. In the context of the Famine, individuals residing in severely affected regions who survived through the support of their clan culture, would have their trust in clan members reinforced by this experience.

1.2.2 Clan as a Kinship-Based Network and Risk Sharing Institution

Clan is a kinship-based organization that includes patrilineal households with a shared lineage tracing back to a common male ancestor. Similar to the corporation, a voluntary organization between unrelated individuals, clan sustains cooperation among members and provides local public goods (69). However, the nature of cooperation within a clan is grounded in reciprocal moral obligations and communal moral values, regulated through the kinship network. In contrast, cooperation within a corporation is based on generalized moral obligations regardless kinship (69; 51)

The most famous metaphor for clan is presented by by (55) " kinship - is similar to concentric circles formed when a stone is thrown into a lake...every family regards its own household as the center and draws a circle around it. This circle is the neighborhood, which is established to facilitate reciprocation in daily life...This pattern of organization in Chinese traditional society has the special quality of elasticity" (p63-64). This "egocentric" network shown by Fei, is defined as *differential mode of association* (chaxugeju). This kinship-based network lacks clear boundaries. [7] The trust within clan members largely depends on their biological distance. Shown by the upper panel in Figure A1, in a society with stronger clan culture, neighbors and relatives are positioned in a more central circle around the individual, reflecting a higher level of trust in them. In contrast, in a society with a weaker clan culture (lower panel), neighbors and relatives are situated in a circle closer to strangers, indicating lower levels of trust in them. One clarification is necessary: given our focus on rural residents in China, it's important to note that neighbors are often relatives. Therefore, in this paper, we do not distinguish between neighbors and family members.

Within the kinship-based network , clan members tend to promote codes of good conduct and supply communal goods (4). Clan provides local militias during turbulence (116). Additionally, clan plays a vital role by providing charity, informal lending, and mutual insurance, serving as institutions for resource pooling and risk-sharing (39). This, in turn, helps reduce survival risks during famines and wars, thereby boosting population growth (38).

A unique institutional feature of clan is the use of genealogy book, which detailedly record the family tree. These books serve as a vital link connecting all males, both past and present, within the kinship network. They play a crucial role in determining the membership of each household within the clan (29). Moreover, genealogies document the codification of clan rules, establishing and reinforcing reciprocal moral obligations among clan members (39). Along with ancestral hall, which serves as a physical space for ancestor worship and important events, genealogy serves as a pivotal tool to establish group identity among clan members. The compilation and upkeep of genealogies demand a substantial economic investment and a high level of cooperation within the clan. Therefore, we posit that the density of genealogies serves as a systematic proxy for the strength of clan culture.

[7] In contrast, individuals in a Western-style organization typically enroll or sign up for memberships.

1.2.3 Clan during the Famine

The unprecedented nationwide famine during the Great Leap Forward (GLF) movement resulted in 30 million deaths from 1959 to 1961. The inflexible excessive procurement by the upper-level government and the misallocation of food resources were considered as the main causes of the Famine (99). During the GLF, a compulsory grain procurement system was initiated, and private ownership of grain was prohibited. Each year, prior to the harvest, the procurement quota is determined, taking into account reported grain outputs from previous years, weather conditions, and historical grain suitability (99; 87). Counties with favorable weather conditions and a historical suitability for grain crops are associated with higher procurement quotas. As a result of career and promotion incentives, upper level Communist Party officials over-reported grain outputs, leaving insufficient food crops for local communities to sustain themselves. This, in turn, was one of the contributing factors to the occurrence of the Famine. As documented by (130), at the peak of the GLF, rural households were no longer permitted to store their own food.

Under pressure from upper-level officials, village leaders from regions with higher clan strength were more likely to resist excessive procurement or conceal grain from the upper-level government. Meanwhile, inter-clan borrowing served as a channel to save their clan members and alleviate the intensity of the Famine. As a result, (29) finds that the increase in mortality rates during the Famine years is significantly smaller in counties with a higher level of clan strength. Particularly, a one standard deviation increase in clan strength is associated with a reduction of 1.71 to 2.26 deaths during the Famine.

1.3 Data

This section discuss our main data sources and key measurement strategies. Our empirical strategy make use of four main data sources: Trust outcomes from *China Family Panel Study* (CFPS) at individual level, famine intensity from county gazetteers, clan strength from historical collection, and historical weather shocks and soil suitability index sourced from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) V4.0 database. More details about the summary statistics can be found in table 1.

1.3.1 Individual Level Trust

We measure the main outcome of interest, trust in neighbors, using the second wave of the *China Family Panel Survey* (CFPS) survey (2012) and gather individual characteristics from the baseline survey (2010). CFPS is a nationally representative longitudinal survey launched in China in 2010, focusing on Chinese communities, families, and individuals. It is regarded as the counterpart to the PSID data in the United States. Shown in Table 1, after matching our data with famine intensity and clan strength, we retained 92 counties and 7514 individuals for our main analysis.

The main outcome of interest comes from the question:

To what extent do you trust your neighbors?

(where 0 means that you have complete distrust and 10 means that you have complete trust.)

Shown by Table 1, the average trust in neighbors is 6.5 out of 10 points. In contrast, trust in parents has an average score of 8.9, while trust in strangers averages only 2. This pattern aligns with common sense. Regarding generalized trust, a widely-used measure of social capital (e.g., in the World Values Survey and General Social Survey), half of the respondents in our sample agree that "most people can be trusted". For purpose of robustness, we also create dummy indicators to classify the continuous trust measures as above or below 6 points.

For our baseline estimation, we restrict the sample to individuals born between 1941 and 1970, who resided in rural areas and lived in the same counties since their birth. We also use urban counterparts as a falsification test.

1.3.2 County Level Famine Intensity

We calculate county-level famine intensity using data from county gazetteers, government reports, and population statistics compilations. Our main data source is from *China Gazetteer Project* [8], a large scale project to digitize local gazetteers at Harvard's Yenching Library. County gazetteers are local encyclopedia covering major events since 1949 to 1990s, including democratic information, economic development, political movements, agricultural production and so on [9].

In particular, We collect annual death counts per thousand people for each county and match them with CFPS sample counties. Then, we define the death rate in famine years as the average death rate during 1959 to 1961 and the death rate in normal years as the average death rate during 1954 to 1957 [10]. Finally, we use county-level excess mortality as a proxy for famine intensity, calculated as the ratio of the death rate during famine years to that during normal years, minus 1. Our measure of famine intensity can be interpreted as the percentage increase in deaths during famine years compared to normal years. This measure addresses concerns related to differential death rates caused by varying age profiles across counties. Figure A5 displays a fat-tailed distribution of famine intensity for CFPS sample counties. The sample mean is 0.89 and sample median is 0.43. In our baseline regression, we use a dummy indicator based on whether the excess mortality level is above or below the sample median.

It is reasonable to consider that the mortality data compiled in county gazetteers and government statistics may be under-reported. However, most of the data we utilize were compiled in the early years of the reform (1980s), when the local officials responsible for famine deaths were no longer in office, and people began to reevaluate the disasters during the Great Leap Forward and the Cultural Revolution. During our interview with one of

[8] https://chinagazetteer.wixsite.com/project

[9] A growing number of literature in Chinese study use this data source, see (36), (35), (29)...

[10] We exclude the years 1949 to 1953 from the normal years due to the ongoing regional civil war and land reform during this period. 1953 is considered the first year of large-scale economic construction. Additionally, we exclude the year 1958 from normal years because historical evidence indicates that the famine had already begun in some counties during that year.

the county gazetteer editors, he assured us that all the statistics included in the gazetteer are accurate, as they were required by higher-level officials [11]. As part of our robustness analysis, we additionally employ a cohort loss index, calculated based on relative cohort size, as an alternative proxy for famine intensity. (99).

1.3.3 County Level Clan Strength

As previously discussed in Section 2.2, the compilation and maintenance of a genealogy require dedication from clan members. The existence of these genealogy books serves as a proxy of clan strength and social capital. We use the density of genealogies as our main measure of clan strength, a similar measure employed by (29; 39; 68). Specifically, we collected geographic information of 30330 genealogy compiled before 1950, sourced from *The General Catalog of Chinese Genealogy*, recognized as the most comprehensive registry of Chinese clan genealogies to date.((69); (48)) [12]. Considering that clan members typically reside in close-knit, compact communities (38), our underlying assumption for measuring clan strength at the county level is that counties with higher genealogy density indicate a larger proportion of communities within that county being associated with clans.

We first take the logarithm of the per capita count of genealogies compiled before 1950 in a county, which is normalized by the population recorded in the 1953 census. Then, we generate a dummy variable that indicates whether the clan strength is above or below the sample mean [13]. Table 1 shows 26% of CFPS samples counties are categorized as as having high clan strength. Figure A3 displays a national geographic distribution of log genealogies per capita compiled before 1950. White regions represent counties where no genealogy books were compiled during the investigated time span. We observe that the distribution of genealogy books is concentrated in the southeastern regions of China, aligning with the historical narrative of clan distribution.

[11] Check our interview with Hu Erson, the editor of the Pingu Gazetteer. https://www.youtube.com/watch?v=_NUY39R31s8

[12] This data set is digitized by (133) and public available now.

[13] the sample mean of log (normalized_genealogies) is 0.13, median is 0.

Several important caveats should be discussed. First, there might be survivorship bias, as some genealogy books may have been destroyed before the publication of the Catalog. However, this bias could potentially strengthen the proxy for clan strength, as genealogy books are more likely to survive in counties with strong clan adherence. Second, the land reforms in the 1950s weakened or eliminated many local landowning families. As a result, clan strength before 1950 could not predict the clan strength during the Famine (1959 to 1961). we will check how many clans survived the land reforms until late 1950s.

1.3.4 County Level Soil Suitability and Weather Shock

The data regarding the suitability of soil for different crops are sourced from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) V4.0 database. This database offers detailed information on the potential yields of various crops under various technologies at a grid level of the 9.25km $\times$ 9.25km. To accurately capture the farming technologies used in China during the 1950s and 1960s, we adopt the methods provided by (99; 98). We select the production function considering rain-fed irrigation, intermediate input levels, and no CO2 fertilization. This suitability measure serves as a time-invariant index, reflecting the suitability of regions for cultivating key procurement crops in China during the 1950s, including rice, sorghum, wheat, buckwheat, and barley. We calculate the soil suitability index at the county-level by averaging the values of the grids within each county's boundaries across all selected crops [14].

The historical weather data are sourced from the China Catchment Attributes and Meteorology dataset (CCAM), which offers daily temperature and precipitation records at the meteorological station level (75). First, we calculate the county-level daily weather variables using Inverse Distance Weighted (IDW) interpolation with data from the five nearest meteorological stations. Then, following the methods of (99; 87), we create variables for average temperature and precipitation during the Spring months (February, March, and

[14] In the V4.0 database, there are only options of high and low inputs. However, according to the document of V3.0, the intermediate input is just the average of high and low inputs.

April) and Summer months (May, June, and July) for each county-daily observation. Lastly, similar to our approach with excess mortality, we define weather shocks as the percentage deviation in temperature and precipitation during famine years compared to normal years. [15].

Since the weather and soil suitability data were collected for scientific research purposes, there is no evidence to suggest that the Mao-era government manipulated the data

1.4 Empirical Strategies and Results

In this section, we present two primary empirical strategies and results. Firstly, employing a difference-in-differences estimation, we demonstrate that exposure to the Famine resulted in a statistically insignificant, slightly negative, impact on trust in neighbors. Nonetheless, this traumatic experience increased trust among individuals residing in counties with high initial social capital. Secondly, through a triple-differences strategy, we offer additional evidence of the heterogeneous response to this traumatic experience across different initial conditions.

1.4.1 Difference in Differences Estimation

In the first part of the analysis, we study the impact of the Famine on the entire sample and as well as on subgroups categorized by their initial clan strength. We limit our CFPS sample to individuals born between 1941 and 1970, who resided in rural areas and lived in the same counties since their birth. In particular, we exploit variations in county-level famine intensity exposure among cohorts before and after the Famine in a difference-in-differences (DiD) setting:

[15] please check Appendix X for detailed data cleaning process.

$$T_{ivct,s} = \beta_1 \times Mortality_{c,s} \times Cohort_{t,s} + \Gamma \mathbf{X}_{ivct,s} + \gamma_{v,s} + \gamma_{t,s} + \varepsilon_{ivct,s} \tag{1.1}$$

where the subscripts i denotes a individual, v denotes community, c denotes county, t denotes the year of cohort birth and s denotes the samples used for analysis, including the entire sample, the high clan strength group, and the low clan strength group. T_{ivct} is individual level of trust in neighbors. $Mortality_{c,s}$ is a dummy variable takes the value of 1 when county-level excess mortality during the Famine exceeds the sample median (0.43), as explained in Section 3.2. $Cohort_t$ is a dummy variable for whether that individual was born before the Famine (1961). $\mathbf{X}_{ivct,s}$ contains individual controls including gender, education and ethnicity. $\gamma_{v,s}$ are the community (village) level fixed effects, and capture time invariant characteristics across villages. $\gamma_{t,s}$ are cohorts fixed effects common to all individuals in $Cohort_{t,s}$. $\varepsilon_{ict,s}$ is idiosyncratic errors. Rousted standard errors clustered at county level.

According the conceptual framework, we examine whether β_1 is significantly positive in the DD specification for the high clan strength group. The low clan strength group serves as a placebo test. We should not expect to observe an effect on trust in their neighbors, as there was no strong initial social capital during the Famine.

1.4.2 Difference in Differences Results

The estimates from equation (1) are in Panel A of Table 2. As the Column 1 shows, the estimated coefficient is negative and statistically insignificant for entire sample. the estimated coefficient is negative and statistically insignificant for the entire sample. This finding aligns with the varied results from previous studies examining the impact of traumatic experiences on social capital. In Column 2, the results for individuals residing in counties with high initial clan strength are consistent with the conceptual framework.

Specifically, the strong famine-exposed cohort experienced an increase in trust in their neighbors after surviving the Famine. The increase in the trust score by 0.75 is quite substantial, particularly when compared to the sample mean of trust scores in neighbors, which stands at 6.5 points. In Column 3, as demonstrated for individuals living in counties with low initial clan strength, the effect is relatively modest and statistically insignificant. This suggests that there was no significant update in their beliefs regarding social capital, since they did not benefit from this risk-sharing institution. The F-test between the estimators derived from these two subgroups yields an F-statistic of 8.878 with a p-value of 0.0037. This implies that the results exhibit statistically significant differences in initial clan strength.

To address the concern that our findings might be sensitive to the choice of the cutoff used to generate the mortality dummy, we also employ a continuous measure of excess mortality as treatment, as illustrated in Panel B of Table 2. Our findings concerning remain consistent. In particularly, a 10-percentage-point increase in famine intensity raises the trust score by 0.046 points for the high clan strength group. Additionally, we notice a negative, though smaller, effect on the low clan group.

1.4.3 Event Study— DiD

The parallel-trends assumption is crucial to our analysis. We plot event study graph versions of equation (1) as following:

$$T_{ivct,s} = \sum_{t=1}^{12} \beta_{1,t} \times Mortality_{c,s} \times Cohort_{t,s} + \Gamma \mathbf{X}_{ivct,s} + \gamma_{v,s} + \gamma_{t,s} + error_{ivct} \tag{1.2}$$

where $Cohort_{t,s}$ is an indicator function denoting whether the individual's birth year falls within a three-year birth cohort bin between 1941 and 1977, originating from either high or low clan strength counties [16]. The cohort born between 1962 and 1964 serves as the

[16] There are two underlying reasons behind this choice. Firstly, the Famine spanned three years, so it is

reference group. We expect that the coefficient $\beta_{1,t}$ will be statistically indistinguishable from zero for cohorts born after the Famine. However, for cohorts within high clan groups that have reached an age to have experienced the Famine, we anticipate β_1 to be positive. To allow for more post-famine periods, we also include individuals born between 1970 and 1977 in the CFPS in the data used for estimation.

Figure 1 plots the event-study estimates for the famine effect by subgroups, with the x-axis plotting three-year birth cohort bin. In the upper panel, which represents the high clan strength sample, we can observe that the positive famine effect is not driven by post-trends, as there is no significant impact on the outcomes of individuals born after the Famine. The famine effects are notably strong for cohorts born between 1953 and 1961, who were aged 0 to 9 during the Famine. In the lower panel, which represents the low clan strength sample, we observe null effects across all cohorts. This observation aligns with the conceptual framework, indicating that individuals live in low initial social capital counties do not update their beliefs in clan network following the traumatic experience shock.

1.4.4 Triple Differences

To estimate how the Famine's impact on trust in neighbors varies with initial clan strength, we employ a triple differences strategy that exploits three sources of variation. Initially, we utilize the variation in famine intensity at the county level combined with the cohort variation in the DiD estimation discussed in Section 4.2. Additionally, we exploit the county-level variation in initial clan strength to estimate the differential famine effects across various initial conditions. These combined sources of variation in famine intensity, cohort, and initial clan strength results in the triple-differences strategy as following:

logical to establish a single group for those who partially experienced the Famine (born between 1959 and 1961). Secondly, we group cohorts into three-year bins to increase the statistical power.

$$
\begin{aligned}
T_{ivct} = \beta_1 \times Mortality_{c,v} \times Cohort_t \times HighClan_{c,v} + \beta_2 \times Mortality_{c,v} \times Cohort_t \\
+ \beta_3 \times Cohort_t \times HighClan_{c,v} + \Gamma \mathbf{X}_{ivct,s} + \gamma_{v,s} + \gamma_{t,s} + \varepsilon_{ivct,s}
\end{aligned} \tag{1.3}
$$

Where $HighClan_{c,v}$ is a dummy variable indicating whether the clan strength, as measured by the logarithmized historical genealogy books per capita, exceeds the national mean. Our coefficient of interest is β_1. We use entire sample for analysis and cluster robust standard error at the county level.

Panel A of table 3 reports the triple-differences estimates. Column (1) to (3) show that individuals exposed to the famine from counties with initially high clan strength exhibited increased trust in their clan members afterward in more famine affected counties. The traumatic experience enforce trust among those who received support from their clans, resulting in an increase in trust scores ranging from 0.42 to 0.66 points. These effects are noteworthy, given that the average trust score in neighbors stands at 6.47. The results remain stable after we control individual characteristics (Column (2)) and allow cohort trends to differ across provinces (Column (3)).

Panel B of table 3 reports the estimates with an alternative specification: Dummy variable $Mortality_{c,v}$ is replaced by a continuous variable representing excess mortality rates, and the same replacement applies to the secondary and triple interactions. We observe that the pattern of results remains stable.

1.4.5 Event Study— Triple Differences

Furthermore, we plot event study graph versions of the triple-differences specification, as outlined in Equation (4):

$$T_{ivct} = \sum_{t=1}^{12} \beta_{1,t} \times Mortality_c \times Cohort_t \times HighClan_c + \sum_{t=1}^{12} \beta_{2,t} \times Mortality_c \times Cohort_t + \sum_{t=1}^{12} \beta_{3,t} \times Cohort_t \times HighClan_c + \Gamma \mathbf{X}_{ivct,s} + \gamma_v + \gamma_t + error_{ict} \tag{1.4}$$

The cohort born between 1962 and 1964 serves as the reference group. We use entire sample for analysis and cluster robust standard error at the county level. We also include individuals born between 1971 to 1977 for more post-famine periods.

In figure 2, we observe that there is no significant post-trends among cohorts who born after the Famine. Among the cohorts that experienced the Famine, the positive effects on trust in neighbors persist for individuals residing in counties with higher clan strength.

1.5 Instrumental Variable Strategy

Although the parallel trends enables us to mitigate potential sources of bias in the estimates, a natural concern is the correlation between famine intensity and initial clan strength. As found by (29), counties with higher clan density significantly reduce mortality during the Famine. It is likely that the famine intensity is correlated with interaction term between the clan strength and cohort dummy. Consequently, the ideal measurement of famine intensity should be exogenous to the initial clan strength. According to (99), the inflexible procurement system is the main cause of the Famine. However, the only available data on procurement is the actual amount of procurement rather than target quotas set before the agricultural season (87). Instead, we use soil suitability and weather shocks during the Famine years as instrumental variables for famine intensity.

1.5.1 Logic of Soil Suitability and Weather Shocks as Instruments

During the famine years, there is a significantly positive correlation between higher production and higher mortality, as noted by (99; 87). Counties with soil more suitable for grain crops and experiencing favorable weather conditions tend to receive higher procurement quotas, which, in turn, lead to more severe famine.

Additionally, soil suitability for crops and weather shocks during the famine years are unlikely to be correlated with initial clan strength. Although (127) find that a history of farming rice promotes cooperative behavior, whereas farming wheat makes cultures more independent, there is no evidence to suggest that the procurement system exhibited a preference for either rice or wheat. As a matter of fact, following (99), the soil suitability index constructed by us including rice, sorghum, wheat, buckwheat, and barley, is exogenous to clan strength before the Famine. Moreover, while long-term weather patterns may be associated with local cooperative behavior (61; 23), it is unlikely that the weather shocks during the famine years are correlated with clan strength. Details on the construction of instrumental variables of soil suitability and weather shocks can be found in Section 3.4.

1.5.2 Instrumental Variable Results

Table 4, Column (1) and (2) show the results for the first stage regression of $Mortality_{c,v} \times Cohort_t$ on the instruments by clan strength. The F-statistics for excluded instruments are 9.66 for the high clan strength sample and 11.1 for the low clan strength sample, respectively. Column (3) and (4) present the results for the first stage regression of $Mortality_c \times Cohort_t \times HighClan_c$ on instruments with similar specification with baseline estimation, with F-statistics of 10.39 and 10.62.

When merging with the instruments, our sample counties decreased from 92 to 82, shown in Table 1. Therefore, we re-estimated the DiD specification with the balanced 82 counties and present the results in Columns (1) and (3) of Table 5. In Column (2), when considering the Instrumental Variable results, the traumatic effect on trust in neighbors for the high clan strength group remains positive and becomes even larger compared to the DiD estimation. In Column (4), the results remain consistent for the low clan strength group as well.

Column (2) of Table 6 provides the IV results in the triple difference specification, with individual controls, community fixed effects and province by cohort fixed effects, robust standard errors are cluster at county level. When comparing the OLS estimates in Table 5 to the IV estimates in Table 6, it is evident that the IV estimates are larger than the OLS estimates.

1.6 Robustness

1.6.1 Falsification Test

As we discussed above, egocentric network is the foundation for belief update from the soil (rural China) (55). Base on this theory, we should should not observe the impacts of the famine experience on trust in neighbor from urban sample. Similarly, we also should not observe any effects on the trust in parents or strangers — people at the right center or absolutely outside the *differential mode of association* (chaxugeju).

We report the regression results on different trusts, separately by rural sample and urban sample based on their *hukou* status. Table 7 column 5 shows the same result discussed in the last section, a positive and significant effect on the trust in neighbors for the rural sample. Comparing to column 3 and 5, we can find that either the trust in people located nearest or farthest to the concentric point (self), the effect is trivial and not significant. The logic behind this phenomena is straightforward, parents are the closest people and will always help their children out during the famine, the strangers are in the quit opposite. To further explore the impacts on social capital (Column 1), we use the binary variable *general trust* as a proxy [17], taking the value of 1 if the respondent believes that "Generally speaking, most people can be trusted." The result is still close to 0 and insignificant, consist with the

[17] This variable is a standard proxy used by social capital literature such as (112), (28) and (111)

trust in stranger case. Column 2, 4, 6 and 8 present none significant results using the urban sample, consist with our story that the trauma experience interacted with clan strength only affected the trust update for rural population. Furthermore, only the connection between the central point (self) and the middle circle (neighbors and relatives) are tightened by this mechanism, not the intra-nuclear family trust and social trust.

1.6.2 Trust Distance Between Circles

One possible concern to our measurement of trust is some unobserved factors might affect the reported score of trust. For example, respondent A might reports 9 points out of 10 for trust in parents and respondent B reports 8. But in reality, B could has more trust in his parents than A does. To address this concern, We take difference between three main trust variables and therefore to differentiate out the idiosyncratic benchmark error. Table 8 shows the regression results for outcomes of trust distance between parents and neighbors, between neighbors and strangers and between parents and strangers. Consist with the previous result, The only 1 % significant effect is on the trust between parents and neighbors (column 5), for the rural sample. Famine-experienced cohorts shows closer trust distance between their parents and neighbors, comparing to the reference group. But the post famine cohorts are not observed with any significant differences. The circle of neighbors get closer to the concentric point — the clan network got strengthen for the treated group. In contrast, The trust distance between neighbor and stranger becomes larger — people who are saved by their clan would be more alienate to the civil society.

1.7 Conclusions

This paper examines the evolution of trust among clan members, in the context of Great Chinese Famine (1959-1961), depending on the historical level of Chinese clan culture. We gather information on clan strength from genealogy books and compile data on famine intensity from county gazetteers. Our triple-differences analysis exploit county-level variation in famine intensity, county-level variation in clan strength and variation in famine exposure base on birth cohort.

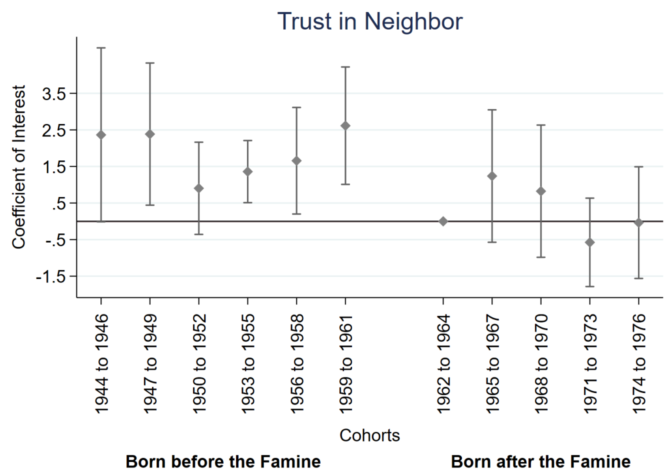
Our analysis shows that the famine exposed cohort that live in a stronger clan county report higher level of trust in their clan members, relative to the people who didn't perceive a sever famine. The magnitudes of effects are non-trivial and consistent to a series of falsification tests, robustness checks and instrumental variable estimations. Our results additionally highlight that the famine effect on both generalized trust and trust in parents does not vary across different initial clan conditions. This lends support to our hypothesis, suggesting that clan culture played a crucial role in protecting individuals during the famine, consequently shaping their beliefs.

1.8 Figures and Tables

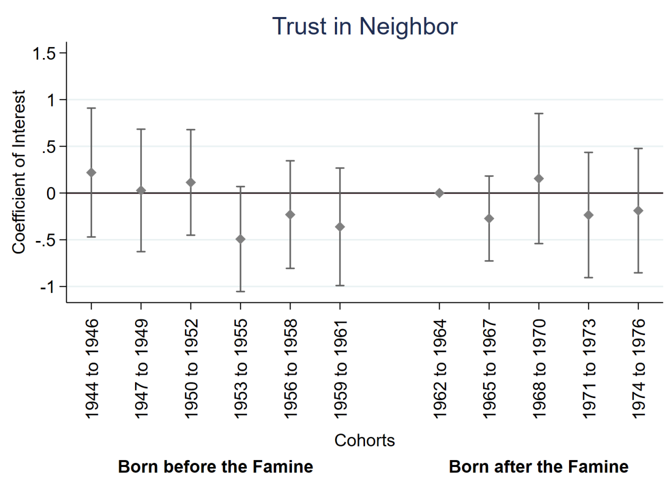
1.8.1 Figures

Figure 1: Event Study for the DiD specification by Clan Strength

(a) High Clan Strength

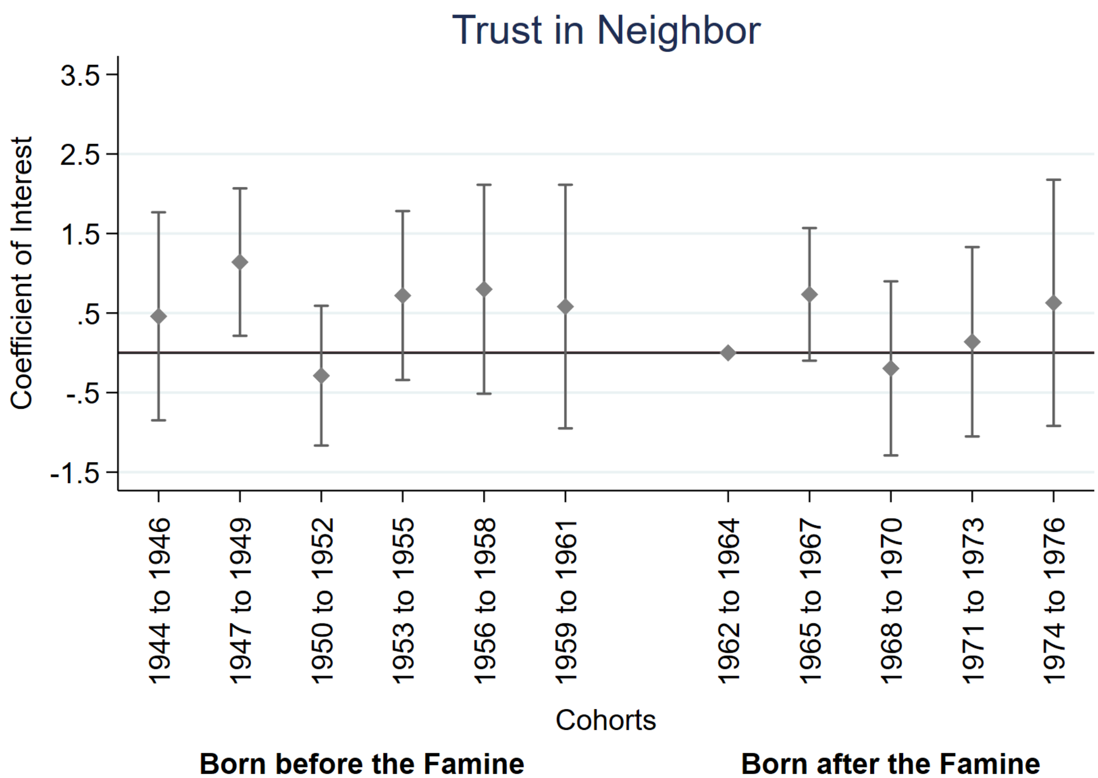


(b) Low Clan Strength



Notes: These figures plot the coefficients for the interaction between being born in a 3-year cohort-group, and the indicator of county excess mortality exceeding the sample median by High / Low clan strength. The cohort from 1962 to 1964, born right after the famine, serves as the reference group. standard errors are clustered at county level.

Figure 2: Event Study for the triple-differences specification



Note: This figure plotthe coefficients for the triple interaction term.

1.8.2 Tables

Table 1: Statistic Summary

	Obs	Mean	Std.Dev	Min	Max
Individual D.V.					
Trust in Parents	7463	8.828	1.835	0.000	10.000
Trust in Neighbor	7514	6.474	2.236	0.000	10.000
Trust in Strangers	7465	2.063	2.107	0.000	10.000
Generalized Trust	7479	0.498	0.500	0.000	1.000
Individual Controls					
Gender	7514	0.534	0.499	0.000	1.000
Ethnic Minority	7504	0.130	0.336	0.000	1.000
Linguistic Minority	7342	0.040	0.197	0.000	1.000
Education Level	7514	2.018	1.014	1.000	6.000
County Level					
Excess Mortality ($\times$100%)	92	0.885	1.317	-0.402	6.125
Mortality (Dummy)	92	0.500	0.503	0.000	1.000
Clan Strength (Genealogy books per capita in log)	92	0.130	0.341	0.000	2.408
High Clan Strength (Dummy)	92	0.261	0.442	0.000	1.000
Spring Precipitation Shock ($\times$100%)	82	0.031	0.371	-0.744	1.221
Summer Precipitation Shock ($\times$100%)	82	-0.022	0.172	-0.294	0.678

Table 2: Difference in Differences by Clan Strength

	Trust in Neighbors		
	(1)	(2)	(3)
	Full Sample	High Clan Strenghth	Low Clan Strenghth
Panel A: Mortality (Dummy)			
$Mortality_c \times Cohort_t$	-0.0744	0.752***	-0.0879
	(0.176)	(0.162)	(0.190)
F-Test (High v.s. Low)		F Statistic is: 8.878	P-value is: .0037
Adj R-squared	0.115	0.0797	0.127
Panel B: Mortality (Continuous)			
$Mortality_c \times Cohort_t$	-0.0924	0.462***	-0.103*
	(0.0582)	(0.0645)	(0.0567)
F-Test (High v.s. Low)		F Statistic is: 18.44	P-value is: 0
Adj R-squared	0.116	0.0797	0.127
Observations	7205	1751	5454
Mean of Outcome	6.446	6.512	6.372
Individual Controls	✓	✓	✓
Community FE	✓	✓	✓
Province-Cohort FE	✓	✓	✓

Notes: This table reports difference-in-differences estimates of the interaction between the dummy variable for whether that individual was born before the Famine, and the indicator of county excess mortality exceeding the sample median. All regressions include individual controls, community (village) fixed effects and province by cohort fixed effects. Column 1 reports the results for the entire rural sample, consisting of individuals born between 1941 and 1970, who never moved to another county. Column 2 reports the results for individuals residing in counties with high initial clan strength. Column 3 reports the results for individuals residing in counties with low initial clan strength. Standard errors are clustered at county level.

Table 3: Triple-Differences Estimation

	(1) Trust in Neighbors	(2) Trust in Neighbors	(3) Trust in Neighbors
Panel A: Mortality (Dummy)			
$Mortality_c \times Cohort_t \times HighClan_c$	0.423*	0.454*	0.664**
	(0.247)	(0.248)	(0.273)
$Mortality_c \times Cohort_t$	-0.123	-0.167	-0.231
	(0.140)	(0.145)	(0.171)
$Cohort_t \times HighClan_c$	-0.188	-0.219	0.0513
	(0.159)	(0.158)	(0.170)
R-squared	0.102	0.103	0.116
Panel B: Mortality (continuous)			
$Mortality_c \times Cohort_t \times HighClan_c$	0.262**	0.272**	0.398***
	(0.131)	(0.131)	(0.0901)
$Mortality_c \times Cohort_t$	-0.0584	-0.0746	-0.106**
	(0.0528)	(0.0550)	(0.0454)
$Cohort_t \times HighClan_c$	-0.177	-0.205	0.0588
	(0.155)	(0.152)	(0.163)
Adj R-squared	0.103	0.103	0.117
Observations	7510	7205	7205
Mean of Outcome	6.474	6.474	6.474
Individual Controls	✗	✓	✓
Community FE	✓	✓	✓
Cohort FE	✓	✓	✗
Province-Cohort FE	✗	✗	✓

Notes: This table reports triple-differences estimates of how the Famine's impact on trust in neighbors varies with initial clan strength, exploiting the interaction between the dummy variable for whether that individual was born before the Famine, the indicator of county excess mortality exceeding the sample median, and whether the initial clan strength exceeds the national mean. The sample consists of individuals from rural area, born between 1941 and 1970, who never moved to another county. Standard errors are clustered at county level.

Table 4: IV First Stage

	DID by Clan Strenghth		Tripple Differences (Whole Sample)	
	(1)	(2)	(3)	(4)
	$Mortality_c \times Cohort_t$ (High Clan)	$Mortality_c \times Cohort_t$ (Low Clan)	$Mortality_c \times Cohort_t \times HighClan_c$	$Mortality_c \times Cohort_t \times HighClan_c$
$Suitability_c \times Cohort_t$	0.831**	-0.134	0.00191	-0.135
	(0.369)	(0.153)	(0.00162)	(0.154)
$Precip(Summer)_c \times Cohort_t$	-0.735*	0.471***	0.00264	0.474***
	(0.434)	(0.173)	(0.00213)	(0.173)
$Precip(Spring)_c \times Cohort_t$	1.153***	-0.0946	0.000490	-0.0949
	(0.218)	(0.118)	(0.00130)	(0.119)
$Temp(Summer)_c \times Cohort_t$	-12.36***	1.973	0.0258	2.000
	(4.131)	(2.721)	(0.0399)	(2.737)
$Temp(Spring)_c \times Cohort_t$	0.226	-0.00958***	0.0000374	-0.00948***
	(0.149)	(0.00160)	(0.0000417)	(0.00161)
$Suitability_c \times Cohort_t \times HighClan_c$			0.844**	0.954**
			(0.380)	(0.399)
$Precip(Summer)_c \times Cohort_t \times HighClan_c$			-0.872*	-1.249***
			(0.459)	(0.461)
$Precip(Spring)_c \times Cohort_t \times HighClan_c$			1.140***	1.246***
			(0.225)	(0.244)
$Temp(Summer)_c \times Cohort_t \times HighClan_c$			-13.03***	-14.66***
			(4.307)	(4.952)
$Temp(Spring)_c \times Cohort_t \times HighClan_c$			0.232	0.231
			(0.157)	(0.149)
$Cohort_t \times HighClan_c$			0.0891	-0.481***
			(0.103)	(0.119)
F-Statistic of excluded instruments	9.664	11.10	10.39	10.62
Observations	1301	5134	6435	6435
Individual Controls	✓	✓	✓	✓
Community FE	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓

Notes: The sample consists of individuals from rural area, born between 1941 and 1970, who never moved to another county, categorized by initial clan strength. Columns reports first-stage OLS regressions. Standard errors are clustered at county level. Suitability is measured as soil suitability of counties for cultivating rice, sorghum, wheat, buckwheat and barley. Precipitation and Temperature are measured as the percentage deviation during famine years compared to normal years.

Table 5: IV for DiD by Clan Strength

	High Clan Strenghth		Low Clan Strenghth	
	(1)	(2)	(3)	(4)
	OLS (Balanced Sample)	IV	OLS (Balanced Sample)	IV
$Mortality_c \times Cohort_t$	0.600***	0.836*	-0.113	-0.922
	(0.0405)	(0.443)	(0.249)	(0.757)
F-Statistic of excluded instruments		9.664		11.10
Adj R-Squared	0.0633	0.0934	0.127	0.0661
Observations	1301	1301	5134	5134
Mean of Outcome	6.492	6.512	6.317	6.372
Individual Controls	✓	✓	✓	✓
Community FE	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓

Notes: The sample consists of individuals from rural area, born between 1941 and 1970, who never moved to another county, categorized by initial clan strength. Coefficients are either OLS or IV estimates where excess mortality is instrumented by soil suitability, precipitation and temperature. The interaction between excess mortality and cohort is instrumented by the corresponding variables and the interaction of cohorts. Standard errors are clustered at county level.

Table 6: IV for triple-differences

	Triple Differences	
	(1)	(2)
	OLS (Balanced Sample)	IV
$Mortality_c \times Cohort_t \times HighClan_c$	0.500*	1.589*
	(0.276)	(0.865)
$Mortality_c \times Cohort_t$	-0.104*	-0.712
	(0.0625)	(0.716)
$Cohort_t \times HighClan_c$	0.0478	-0.730
	(0.217)	(0.469)
F-Statistic of excluded instruments		10.39
Adj R-squared	0.123	0.105
Observations	6435	6435
Mean of Outcome	6.454	6.423
Individual Controls	✓	✓
Community FE	✓	✓
Province-Cohort FE	✓	✓

Notes: The sample consists of individuals from rural area, born between 1941 and 1970, who never moved to another county, categorized by initial clan strength. Coefficients are either OLS or IV estimates where excess mortality is instrumented by soil suitability, precipitation and temperature. The triple interaction and the interaction term involving mortality are instrumented by the corresponding variables and the interaction of cohort or / and initial clan condition. Standard errors are clustered at county level.

Table 7: Triple-Differences on different trust scores

	trust score between parents & stranger		trust score between neighbor & stranger		trust score between parents & neighbor	
	(1)	(2)	(3)	(4)	(5)	(6)
	Rural	Urban	Rural	Urban	Rural	Urban
mortality_prefamine_clan	-0.406	-0.401	0.00565	-0.0176	-0.413*	-0.384
	(0.386)	(0.959)	(0.332)	(0.663)	(0.232)	(0.826)
mortality_prefamine	-0.00316	-0.155	-0.0368	-0.228	0.0353	0.0726
	(0.167)	(0.591)	(0.151)	(0.392)	(0.140)	(0.540)
prefamine_clan	0.347	-0.553	0.175	-1.163**	0.173	0.610
	(0.312)	(0.634)	(0.287)	(0.528)	(0.123)	(0.373)
R-squared	0.140	0.122	0.121	0.120	0.0814	0.0753
Observations	8037	973	8043	973	8038	973
Mean of Outcome	6.820	6.955	4.451	4.487	2.368	2.468
Individual Controls	✓	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Effects on Intra-Clan Relationships

	(1) Visits Relatives	(2) Visits Friends	(3) Visits Relatives/Friends	(4) Neighbor Will Help	(5) Contact Frequency
mortality_prefamine_clan	0.0438	-0.0190	0.0766	0.0532	2.115***
	(0.122)	(0.138)	(0.165)	(0.0364)	(0.595)
mortality_prefamine	-0.0893	0.0207	-0.113	-0.0151	-0.983**
	(0.0608)	(0.0664)	(0.0724)	(0.0186)	(0.389)
prefamine_clan	-0.0273	0.151***	-0.197*	-0.0416**	-0.218
	(0.100)	(0.0517)	(0.115)	(0.0190)	(0.395)
R-squared	0.360	0.276	0.292	0.0653	0.166
Observations	8189	8176	8157	6656	7671
Mean of Outcome	1.684	0.914	0.770	0.924	8.192
Individual Controls	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.9 Appendix

1.9.1 Identification Challenge with Graphic Demonstration

Trust Score		Cohort 1	Cohort 0
Clan	1	8	7
	0	5.5	5

(a) High Mortality Areas

Trust Score		Cohort 1	Cohort 0
Clan	1	7	6.5
	0	6	5.5

(b) Low Mortality Areas

Table : clan strength is orthogonal to contemporaneous political movement

If clan strength is orthogonal to contemporaneous political movement (e.g. Cultural Revolution), the DD estimate from high famine intensity areas is unbiased. Shown in the left table above, The effect from left table equals (8-7) - (5.5-5) = 0.5. The effect from right table equals (7-6.5) - (6 -5.5)=0. The total magnitude will be 0.5-0 =0.5 However, if high clan strength induces high revolutionary intensity (harm trust disproportionately) and impacts cohorts overlapping with our exposed cohorts, the DD estimates is biased. Shown in the table below, the estimate of samples from High mortality areas is 0.3, downward bias from the real effect. Nevertheless, If we adjust the estimate with samples from low mortality areas, Our DDD strategy will give us an unbiased estimate : {(7.6-6.8) - (5.5 -5)} - {(6.6-6.3) - (6-5.5)}=0.5.

Trust Score		Cohort 1	Cohort 0
Clan	1	7.6	6.8
	0	5.5	5

(a) High Mortality Areas

Trust Score		Cohort 1	Cohort 0
Clan	1	6.6	6.3
	0	6	5.5

(b) Low Mortality Areas

Table : clan strength is not orthogonal to contemporaneous political movement

The main challenge to our DDD strategy is cohort-varying county (or lower level) factors that simultaneously affect trust and famine - clan interaction. For example, $\{(7.6\text{-}6.8)$ - $(5.5\ \text{-}5)\}$ - $\{(6.2\text{-}6.1)$ - $(6\text{-}5.5)\}= 0.7$, which is biased.

Figure A1: Trust Circles

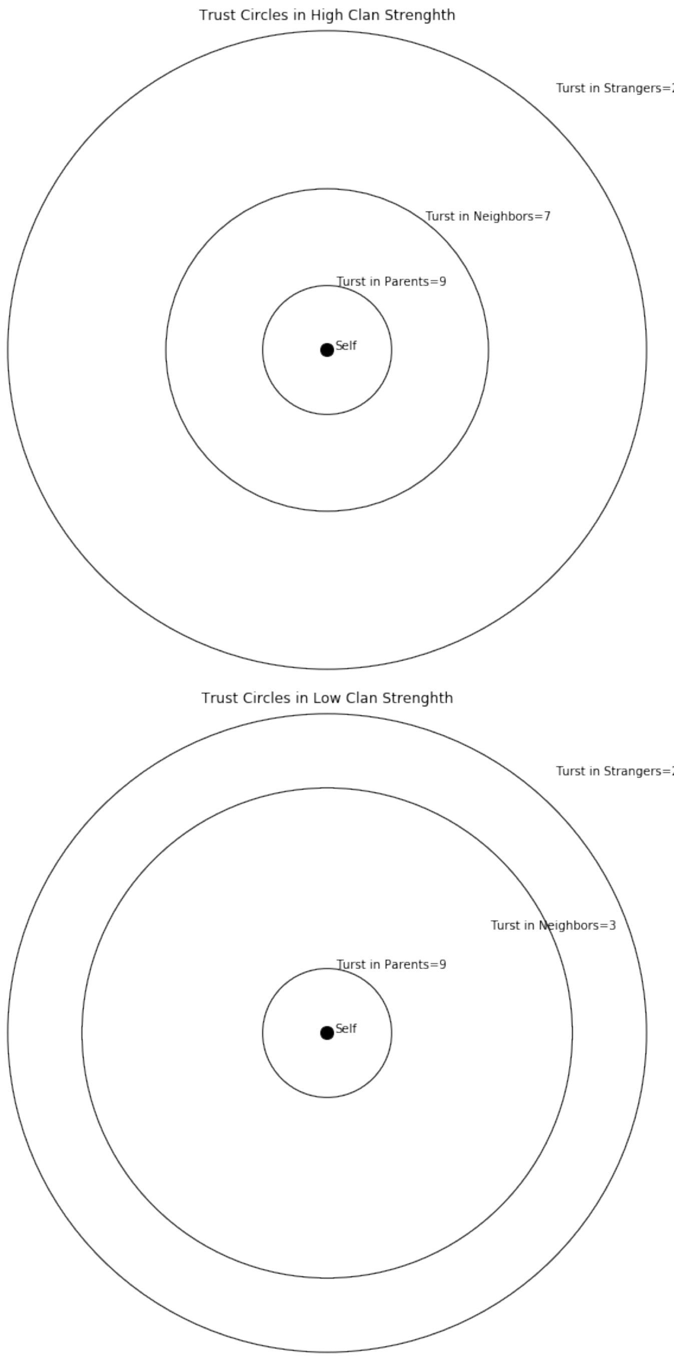


Figure A2: Genealogy Books

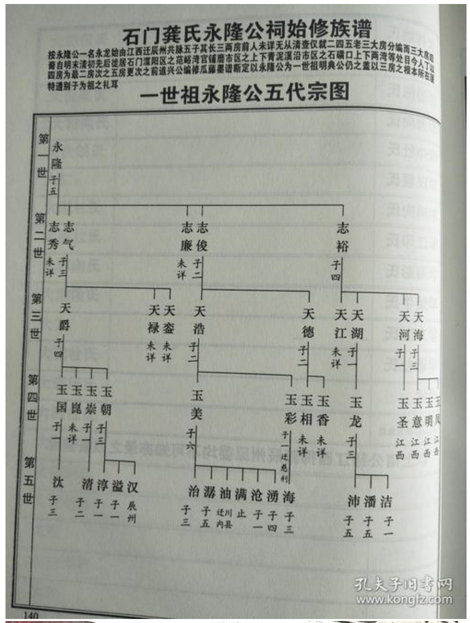
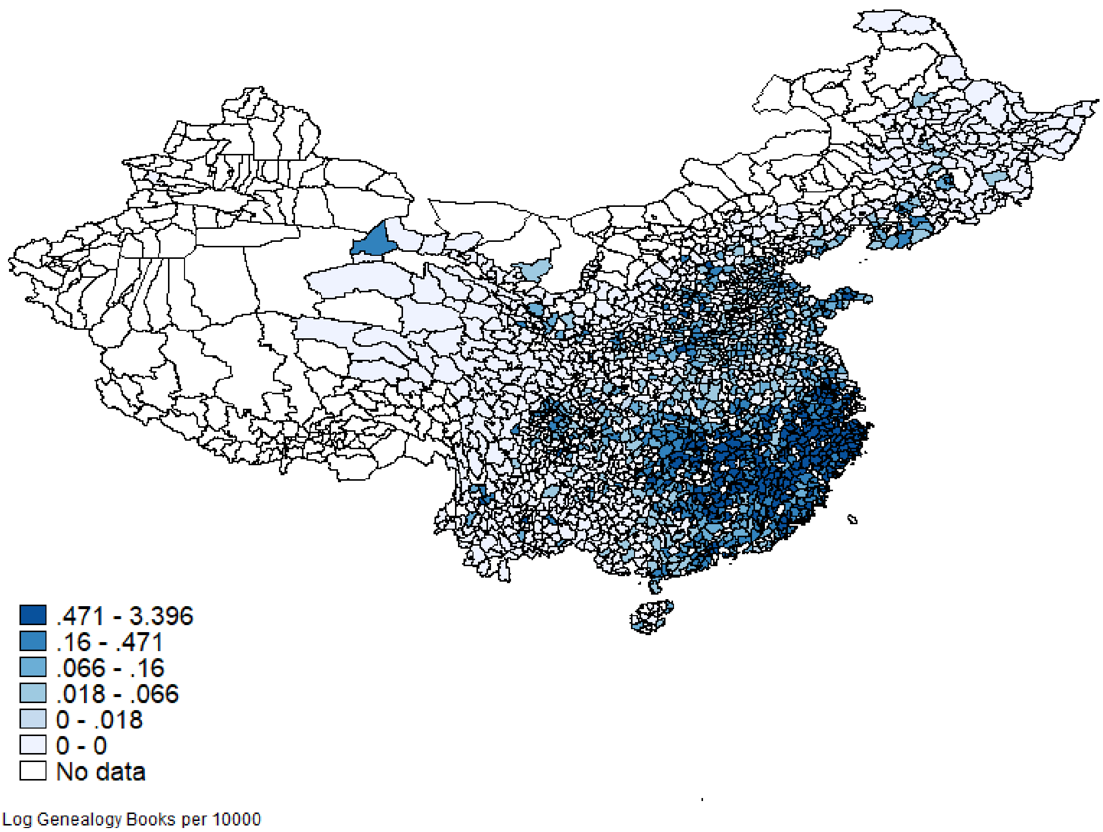


Figure A3: Distribution of Genealogy Books normalized by Population



Note: The county-level clans are measured by the number of genealogy books before 1950 divided by population in 1953, in log form.

Figure A4: Distribution of County-Level Excess Mortality During the Great Chinese Famine

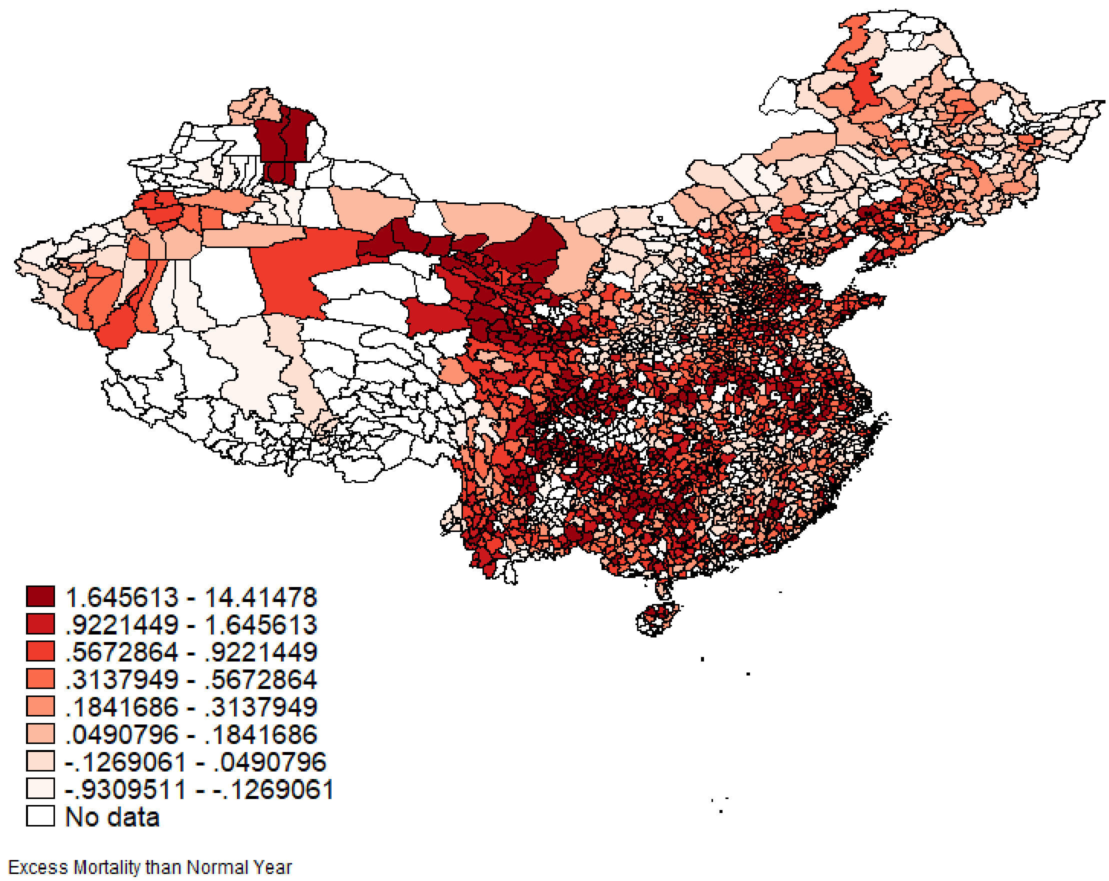
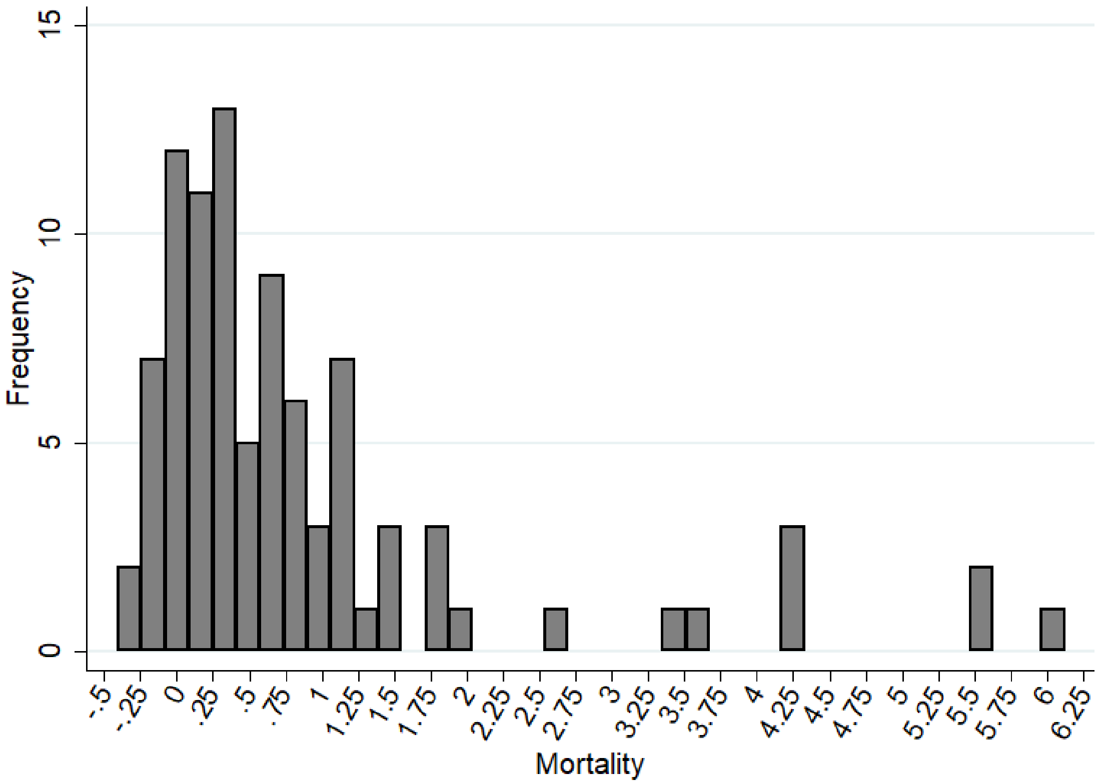
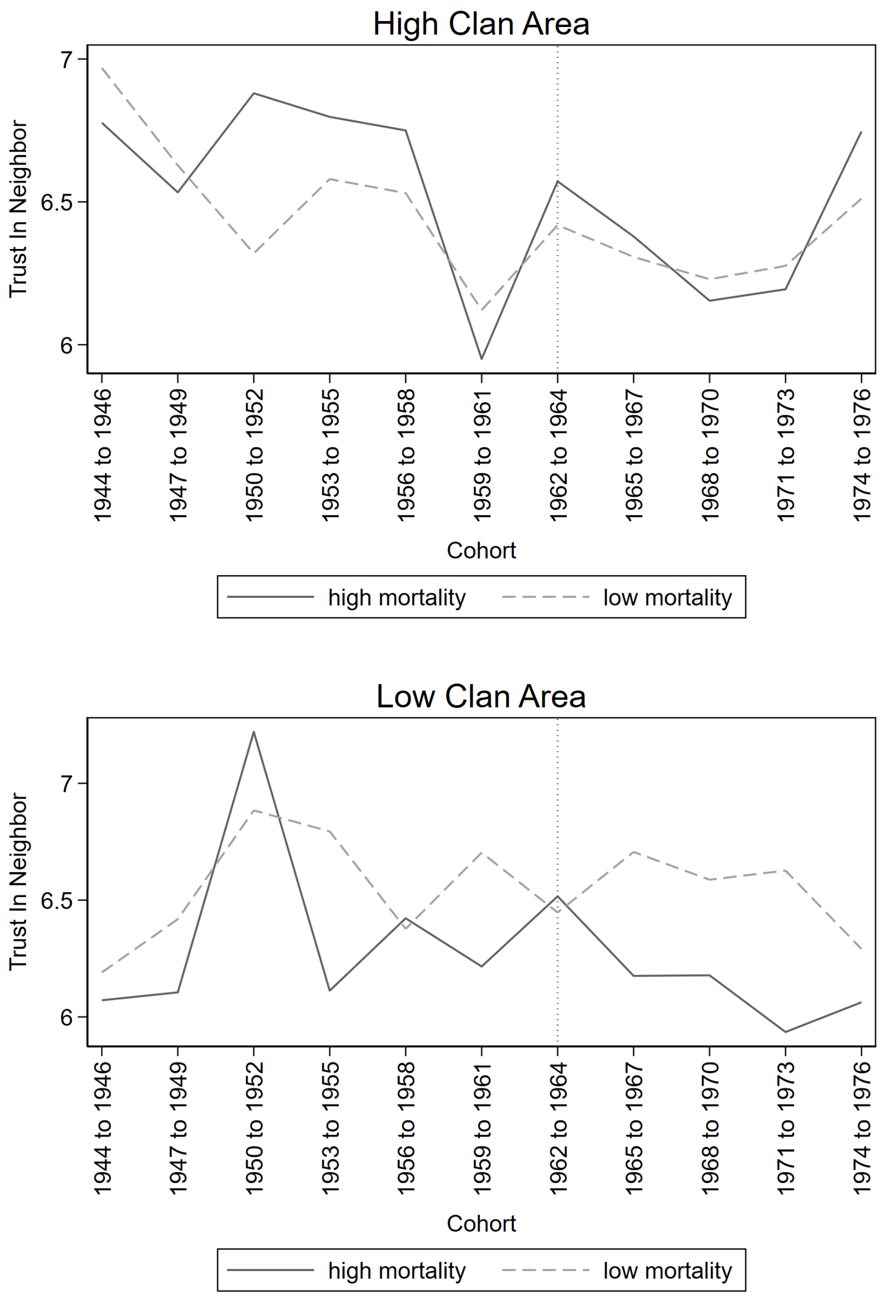


Figure A5: In Sample (CFPS) Mortality Fat Tail



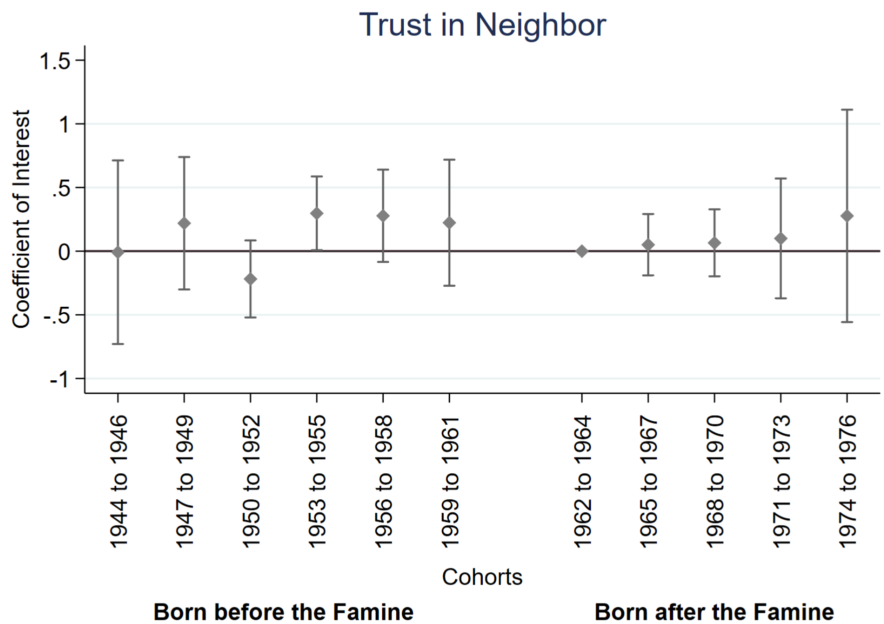
Note: National mean is 0.808, national median is 0.34, in sample mean is 0.89, in sample median is 0.43.

Figure A6: Raw Trust Score across Mortality Levels



Note: The raw score of trust in neighbors by counties of high clan strength and low clan strength for birth cohorts 1941 to 1976, across different mortality levels. Results are based on rural respondents who stay in the origin places.

Figure A7: Dynamic DID Effects by mortality level (in Sample Mean)
(a) High Mortality



(b) Low Mortality

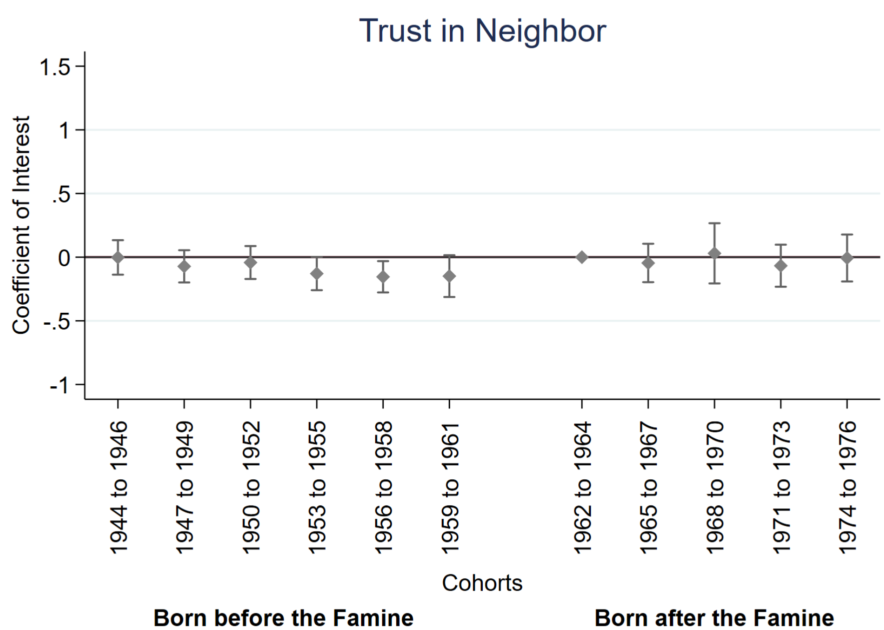


Figure A8: Dynamic Effects of Mortality Dummy (in Sample Mean) on Contemporary Trust

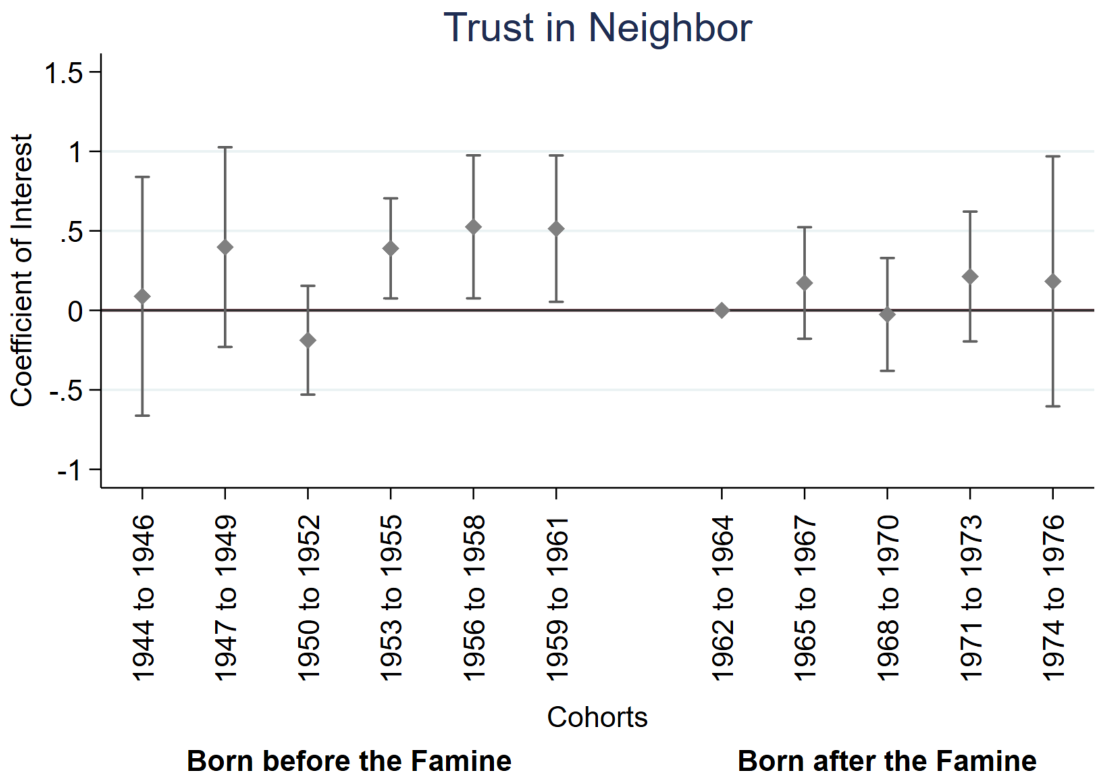


Table A3: Difference in Differences by Mortality Level (in Sample Mean)

	Trust in Parents		Trust in Neighbors		Trust in Stranger		Trust in Neighbors Dummy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	High Mortality	Low Mortality	High Mortality	Low Mortality	High Mortality	Low Mortality	High Mortality	Low Mortality
prefamine_clan	0.361	-0.0602	1.310***	-0.00984	0.00733	-0.139	0.233*	-0.0195
	(0.394)	(0.166)	(0.434)	(0.166)	(0.337)	(0.291)	(0.120)	(0.0288)
R-squared	0.148	0.0916	0.0988	0.101	0.133	0.135	0.104	0.0960
Observations	2312	5634	2315	5638	2315	5637	2315	5638
Mean of Outcome	8.529	9.009	6.262	6.591	2.414	1.881	0.580	0.627
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: The Effects of Famine Experience and Clan Density by Residence (Mortality Dummy in Sample Mean)

	Trust General		Trust in Parents		Trust in Neighbors		Trust in Stranger	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
mortality_prefamine_clan	-0.0293	-0.0408	0.167	-1.985*	1.034***	2.409	0.539	-0.757
	(0.0704)	(0.402)	(0.316)	(1.017)	(0.310)	(1.465)	(0.344)	(0.910)
mortality_prefamine	-0.0178	-0.165	-0.311***	0.234	-0.604***	-1.644	-0.540***	0.222
	(0.0354)	(0.294)	(0.103)	(0.828)	(0.143)	(1.010)	(0.142)	(0.490)
prefamine_clan	0.0366	-0.0459	-0.151	1.223***	0.00718	-0.728	-0.181	0.181
	(0.0490)	(0.257)	(0.154)	(0.423)	(0.156)	(0.902)	(0.281)	(0.891)
R-squared	0.139	0.139	0.131	0.178	0.105	0.180	0.152	0.203
Observations	7992	984	7988	983	7993	984	7992	984
Mean of Outcome	0.488	0.559	8.867	9.273	6.475	6.428	2.024	2.077
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effects on Intra-Clan Relationships (Mortality Dummy in Sample Mean)

	(1) Visits Relatives	(2) Visits Friends	(3) Visits Relatives/Friends	(4) Neighbor Will Help	(5) Contact Frequency
mortality_prefamine_clan	0.144	-0.0518	0.209	0.0858	2.429***
	(0.0897)	(0.258)	(0.243)	(0.0689)	(0.629)
mortality_prefamine	-0.251***	-0.121*	-0.129	-0.0149	-1.208***
	(0.0452)	(0.0686)	(0.0806)	(0.0223)	(0.399)
prefamine_clan	-0.0722	0.137***	-0.224**	-0.0335**	0.160
	(0.0743)	(0.0495)	(0.0854)	(0.0151)	(0.367)
R-squared	0.363	0.277	0.292	0.0656	0.167
Observations	8189	8176	8157	6656	7671
Mean of Outcome	1.669	0.903	0.766	0.921	8.124
Individual Controls	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Household Genealogy Book in 2010 as Clan Measure

	Trust General		Trust in Parents		Trust in Neighbors		Trust in Stranger	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
mortality_prefamine_clan	0.0128	0.171	-0.100	1.739	0.443*	1.769	-0.0523	0.145
	(0.0568)	(0.275)	(0.268)	(1.389)	(0.256)	(1.064)	(0.262)	(1.204)
mortality_prefamine	-0.0208	0.0807	-0.308**	-1.233	-0.381**	0.847	-0.0801	-0.294
	(0.0372)	(0.210)	(0.141)	(0.999)	(0.184)	(0.738)	(0.202)	(0.879)
prefamine_clan	-0.0193	-0.00200	-0.105	-0.645	-0.277	-0.332	-0.0746	0.971
	(0.0324)	(0.135)	(0.163)	(0.858)	(0.198)	(0.714)	(0.156)	(0.774)
R-squared	0.139	0.207	0.130	0.235	0.101	0.169	0.148	0.219
Observations	8371	904	8368	904	8373	904	8372	904
Mean of Outcome	0.492	0.556	8.864	9.138	6.496	6.662	2.042	2.141
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Trimmed Sample (super negative counties removed)

	(1)	(2)	(3)	(4)
	trust_neighbor	trust_neighbor	trust_neighbor	trust_neighbor_dummy
Panel A: Mortality_Dummy Mean (in Sample)				
mortality_prefamine_clan	0.674***	0.650**	0.968***	0.202**
	(0.249)	(0.249)	(0.318)	(0.0852)
mortality_prefamine	-0.235*	-0.229	-0.561***	-0.124***
	(0.133)	(0.139)	(0.148)	(0.0317)
prefamine_clan	-0.226*	-0.218*	0.0522	-0.0201
	(0.132)	(0.130)	(0.164)	(0.0302)
R-squared	0.0921	0.0945	0.104	0.101
Panel B: Mortality_Dummy Median (in Sample)				
mortality_prefamine_clan	0.384*	0.389*	0.465	0.0689
	(0.229)	(0.225)	(0.312)	(0.0746)
mortality_prefamine	-0.177	-0.181	-0.329*	-0.0639
	(0.127)	(0.129)	(0.169)	(0.0428)
prefamine_clan	-0.245	-0.245	0.116	0.00793
	(0.152)	(0.154)	(0.183)	(0.0448)
R-squared	0.0916	0.0941	0.103	0.100
Observations	7614	7609	7563	7563
Mean of Outcome	6.489	6.489	6.489	0.608
Individual Controls	✗	✓	✓	✓
Community FE	✓	✓	✓	✓
Cohort FE	✓	✓	✗	✗
Province-Cohort FE	✗	✗	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Hukou as Rural Residence Measure

	Trust General		Trust in Parents		Trust in Neighbors		Trust in Stranger	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
mortality_prefamine_clan	0.0940	-0.515	-0.178	-0.623	0.671**	2.581	0.816*	1.696
	(0.0694)	(0.435)	(0.281)	(1.377)	(0.321)	(1.654)	(0.451)	(2.185)
mortality_prefamine	-0.0414	-0.0994	-0.224*	1.570***	-0.326*	-0.933	-0.320**	-0.163
	(0.0293)	(0.239)	(0.129)	(0.413)	(0.191)	(0.860)	(0.149)	(1.168)
prefamine_clan	-0.00879	0.0266	0.00875	-0.112	-0.0860	-0.700	-0.314	0.525
	(0.0507)	(0.246)	(0.166)	(0.469)	(0.185)	(1.412)	(0.352)	(1.341)
R-squared	0.135	0.106	0.128	0.135	0.0949	0.204	0.141	0.151
Observations	8639	655	8634	655	8640	655	8639	655
Mean of Outcome	0.491	0.597	8.906	9.153	6.525	6.585	2.030	2.155
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Community FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2

Economic Impacts of China's Zero-COVID Policies

2.1 Introduction

The COVID-19 pandemic severely disrupted general economic activity as human mobility was restricted, social gatherings were banned, and businesses were halted. However, research that examines the effects of the pandemic on the economy has focused primarily on specific areas, such as unemployment, consumer spending, labor demand, and pollution. There is a demand for a comprehensive assessment of the economic consequences of the pandemic and the corresponding anti-contagion policies. Additionally, most of the research has focused only on the year 2020 and has not considered the subsequent periods 2021 and 2022. Our paper aims to fill this gap.

In this paper, we compile a unique dataset of China's COVID-19 risk level on prefecture/county level, which is constructed based on big data provided by the State Council of the People's Republic of China (PRC). We examine the impact associated with China's COVID-19 policies on several salient economic indicators from 2020 to 2022. Specifically, we analyze the effects on mobility, air pollution measured by the concentration of fine particulate matter (PM2.5) and night lights. We rely on a difference-in-differences framework for identification, with the assumption that, conditional on daily confirmed COVID-19 cases and other prefecture-day level controls, the difference in economic indicators between regions with and without COVID-19 containment policies would remain stable over time.

From February 17, 2020, after one month of the pandemic outbreak and a series of strict lockdown measures, China has utilized big data and established a nationwide risk-level system, which aimed to contain the spread of the virus within communities while keeping the economic costs to a minimum, also referred to as "zero-COVID" policy. To be specific, China implemented a nationwide risk response system that mandated local officials to classify communities into low-, medium-, and high-risk levels based on recent confirmed COVID-19 cases and other factors. Areas rated as medium- and high-risk imposed more stringent containment measures compared to low-risk areas, such as stay-at-home order, mass testing, contact tracing and mobility restrictions. Therefore, the classification of an area as *risk* or non-risk is closely linked to the stringency of the zero-COVID policies enforced by local authorities.

It is important to evaluate the economic consequences of zero-COVID policy in the context of both economics and politics. Zero-COVID policies are considered as the Chinese government's pilot experiment in using big data for national management and crisis response.[1] In 2021, China's media outlets portrayed the low mortality rate from COVID-19 as the success of this risk-level system. Moreover, China's GDP growth rate reached 8.1% in 2021. The Chinese government has been promoting their zero-COVID policies as a model for the rest of the world to follow, claiming that it has been effective in both preserving lives while maintaining economic growth. However, in 2022, the emergence of the Omicron variant resulted in shutdowns of financial, manufacturing, and exporting centers, including Shanghai, Shenzhen, Guangzhou, and Changchun, leading to the failure of China's zero-COVID policy to safeguard people's lives and economic vitality (95).

[1] Check out the coverage provided by state-controlled media: https://www.tsinghua.edu.cn/info/1182/51343.htm

Using an original daily panel data at the prefecture/county-level on COVID-19 risk levels collected from the website of the State Council, our study firstly shows that on average the zero-COVID policy took 21 days to eliminate local COVID-19 cases in 2021, but it took approximately 50 days in 2022. Our second finding reveals a 30% reduction in inter-prefecture traffic flow after a prefecture has been classified as a *Risk* region in either 2021 or 2022. Furthermore, our study revealed that the probability of being classified as a *Risk* region was positively and significantly associated with changes in PM2.5 and night lights in 2021, while the effects of the zero-COVID policy are negligible. However, in 2022, the zero-COVID policy led to a decrease in PM2.5 concentration by 1.17% and a reduction in night lights by 7.7%. The differences in policy effects observed between 2021 and 2022 can be primarily attributed to differences in the stringency of the zero-COVID policy. In 2022, with the emergence of the Omicron variant and stricter zero-COVID policies, the negative policy effects on economic activities became significantly larger. Our back-of-the-envelope calculations indicate that the zero-COVID policy caused China to experience a reduction of around 3.9% in GDP in 2022.

The previous studies on COVID-19 pandemic in China have two limitations. First, the majority of studies draw their conclusions focusing on lockdown policies in the early stage of 2020 rather than zero-COVID policies in 2021 and 2022.[2] To date, only one paper has estimated the economic impacts using truck flows in 2020 and 2021 (33). However, it is worth noting that the policy object under study in this paper is prefecture-level city lockdown, rather than zero-COVID policy, therefore it could not account for less stringent policies such as restrictions on human mobility, the establishment of body temperature checkpoints, neighborhood sanitization, monitoring of suspected COVID-19 cases, and other anti-contagious measures at the local community level. Second, they primarily focused on the economic consequences of COVID policies from a single aspect. (47; 64; 136) focus on the COVID-19 policies' adverse effects on labor market outcomes such as unemployment, wage, and labor market participation. Using high-frequency

[2]For example, see (52; 77; 54; 94). For a systemic review, see (84)

transaction data, (31) provided evidence that the pandemic has caused a sharp decline in consumption immediately after the COVID outbreak. (54) documented that the human mobility restrictions imposed by Chinese government in the early phase of the pandemic effectively controlled the spread of the virus. Despite the seemingly high economic and social costs, researchers have also shown that the COVID-19 pandemic significantly improved air quality and reduced environmental pollution (77; 20).

This paper makes three primary contributions. First of all, to be best of our knowledge, our paper is the first empirical study that examines the economic impact of the zero-COVID policy spanning from 2020 to 2022. We offer evidence of the heterogeneous outcomes linked to the implementation of the zero-COVID policy during the three-year pandemic. This research provides insight into the efficacy of the zero-COVID strategy in contributing to China's rapid economic recovery in 2021, and also highlights the disruptions caused by the escalating pandemic and the frequent re-imposition of the zero-COVID policy in 2022. Second, we compiled a unique dataset that reflects the stringency of China's zero-COVID policy. Our dataset provides daily risk level indices at the county level in China from April 2021 to December 2022, including 2853 counties and 368 prefecture-level cities. Local governments have implemented various anti-contagion policies based on risk ratings. The granularity of our dataset could provide new insights and serve as a valuable tool for future research in general to better understand the economic consequences of the pandemic and the zero-COVID policies in China. Lastly, our paper contributes to the existing literature with an in-depth analysis of the economic impact of the COVID-19 policies along three dimensions: human mobility, air pollution, and night lights. The three outcomes in our research offer varying insights into economic performance, such as transportation, manufacturing, and service sectors. Furthermore, the inter-prefecture traffic mobility index and PM2.5 can be used as proxies for short-term economic activities, particularly human mobility and factory productions. On the other hand, night lights can be used as proxies for medium-term economic activities.

The remainder of this paper is structured as follows. Section 2 details the policy background and data. Section 3 delineates the identification strategy. Section 4 presents the main results and performs robustness checks. Section 5 concludes.

2.2 Policies and Data

In this section, we cover basic facts and data source. Initially, we outline China's COVID-19 policies, encompassing lockdown and the zero-COVID. Then, we describe the sources of data for mobility, pollution, and night lights. Finally, we describe the control variables, which include daily confirmed cases and weather.

2.2.1 China's COVID-19 Policy — Lockdown (Jan 23 — Feb 16, 2020)

With the initial COVID-19 outbreak in Wuhan in 2020, the Chinese government implemented unprecedented prefecture lockdown to contain the virus. Stringent measures were put in place in the locked-down prefectures, including the prohibition of traffic leaving, the imposition of stay-at-home orders, and the enforcement of quarantine measures. It's worth mentioning that anti-contagion policies were also enforced in prefectures without lockdowns, albeit with less strict measures compared to the locked-down ones. According to (115), by February 16, 2020, more than 250 prefectures had implemented such measures.[3] Starting from February 17, 2020, the Chinese government implemented a policy package to precisely contain COVID-19 transmission at the community level. As a result, the central government no longer recommended prefecture-level lockdowns, as they were considered too detrimental to the economy. The "Lockdown" in this study is defined as China's major COVID-19 policy from January 23 to February 16, 2020. Our data on lockdowns come from (77), who originally collected from Wikipedia, various sources of news media and government announcements.

[3] "In all Chinese cities, the Spring Festival holiday was extended, and people were advised to stay at home when possible, enforce social distancing and maintain good hygiene." (77)

2.2.2 China's COVID-19 Policy — zero-COVID (Feb 17, 2020 — Dec 25, 2022)

Following the one-month-long enforcement of strict lockdowns and nationwide public health interventions, the central government sought to revive the economy and loosen the lockdown measures. (64). On February 17, Prevention Guidance for Novel Coronavirus Pneumonia (version 5) was issued by the State Council and National Health Commission of China.[4] This guidance mandated local governments to classify COVID-19 risk at the community level. Any community that reported COVID-19 cases would be categorized as either a medium- or high-risk zone, and corresponding containment measures and closures would be enforced. However, in principle, low-risk communities should only impose quarantines on individuals traveling from medium- or high-risk areas and should not limit the traveling of residents or economic activities. The objective of this policy is to eradicate COVID-19 transmission at the local level by assigning each community a risk level and implementing corresponding measures. This is commonly known as the zero-COVID policy.

In order to comply with the guidance, starting from March 2020, the State Council of China began to release a national COVID-19 risk level system on a regular basis through its website. This system categorizes communities within the 2853 counties into high-, medium-, or low-risk groups and updates on a daily basis. All zero-COVID policies, including quarantine, closures of public places, travel restrictions, Travel QR Codes, etc., were implemented based on this system.[5] The COVID-19 risk level system is viewed as a pilot experiment in utilizing big data for national management and crisis response.[6] In particular, the risk level is reported by local governments and compiled by National Health

[4]Prevention Guidance for Novel Coronavirus Pneumonia (version 5): http://www.nhc.gov.cn/jkj/s3577/202002/a5d6f7b8c48c451c87dba14889b30147.shtml

[5]Check out the news from State Council's website: http://www.gov.cn/fuwu/2020-03/25/content_5495289.htm

[6]Check out the coverage provided by state-controlled media: https://www.tsinghua.edu.cn/info/1182/51343.htm

Commission of China.[7] The criteria used to designate a community as either a *Risk* or non-risk area are based on the presence of confirmed cases of COVID-19 reported within recent days. It is important to note that local officials have some flexibility to adjust the coverage range of medium- or high-risk areas. In cases of overreaction, neighboring communities without any cases may still be classified as medium- or high- risk.

Our data on risk level information are drawn from *China's COVID-19 Risk Level Dataset*, a newly constructed dataset containing COVID-19 risk level information for communities within the 2853 counties on a daily basis from April 02, 2021 to December 15, 2022, which marks the end of the zero-COVID policies. This information was collected from the State Council's website (see Appendix A for more details). To the best of our knowledge, this is the first dataset to document China's county-level daily implementation of the zero-COVID policy during 2021 and 2022.[8] We define a county as *Risk* region on a given day if it contains at least one community categorized as medium- or high- risk according to the aforementioned criteria. We define a prefecture as *Risk* region on a given day if at least one community within it is categorized as *Risk* area.

Table 1 shows that on average, from April 02, 2021 to December 15, 2022, 74 counties were classified as *Risk* regions on a daily basis. Averagely, each county was classified as *Risk* region for a duration of 16 days by December 15, 2022 (the end of zero-COVID). Figure 1 shows that the aggregate nationwide daily confirmed cases correlates positively with number of counties with *Risk* areas.[9] Furthermore, we have noticed a steep rise in the number of counties categorized as *Risk* regions beginning in July 2022, while the number of confirmed cases experienced a sharp surge starting only after October 2022. These trends suggest that, comparing to 2021, local officials may be more inclined to enforce stricter zero-COVID policies or potentially overreact with their policies in

[7]The term "risk" used in this context is distinct from its traditional usage in economic research, which involves prediction and expectation. Here, "risk" refers to the assessment of COVID-related risk based on the current presence of COVID-19 cases.

[8]The previous research mainly focus on 2020 or lockdowns, rather than 2021 and 2022 or zero-COVID.

[9]Shanghai is excluded from the sample due to a skyrocketed increase in COVID-19 cases during April 2022.

response to the more transmissible Omicron variant in 2022. This finding is further supported by Figure 2, which illustrates a comparison between the green bar and blue bars. The results show that in 2022, there were much more counties classified as *Risk* regions for longer duration compared to 2021. Additionally, Figure A2 indicates that only a small fraction of counties were classified as *Risk* regions in 2021, whereas by the end of 2022, 1700 out of 2853 counties were classified as *Risk* regions.[10]

There are three things worth noting. Firstly, our binary variable of a county classified as *Risk* or non-risk region does not differentiate the level of intensity of treatment. For instance, a county with only one community designated as *Risk* area and another county with 100 communities designated as *Risk* areas are likely to receive varying impacts from zero-COVID policies. Although there will be differences in the treatment, we are unable to distinguish between them. Secondly, our risk level data does not provide information on the specific zero-COVID policies implemented in each county. For example, if two counties with the same number of communities are classified as *Risk* areas, County A may require all residents to stay home, while County B may only quarantine individuals who have tested positive for COVID-19. The bottom line is that as long as a county/prefecture is categorized as *Risk* region, corresponding zero-COVID policies will be implemented in this region. Finally, a prefecture-wide lockdown remains as an option within the zero-COVID policy framework for the years 2021 and 2022,[11] despite variations in official terminology like "citywide static management", "silence period" and so on. Our research does not aim to differentiate between lockdown and other aspects of the zero-COVID policy during 2021 and 2022. Instead, we regard our estimates as capturing the average impact of a range of interventions, including both stringent measures like lockdowns and milder restrictions.

[10]See Panel B of Table 1

[11]Prominent cities such as Xi'an and Shanghai implemented lockdown measures, with Xi'an being in lockdown for approximately a month starting from the end of 2021, and Shanghai undergoing a lockdown for about four months during the first half of 2022.

2.2.3 Mobility

We use the data from the Baidu Qianxi (Migration) website, which is publicly shared by (83), to construct our measures of human mobility. Baidu is the largest search engine in mainland China. Their migration data are based on real-time location records for every smart phone that uses the company's mapping app, and thus can accurately reflect population mobility between cities.

The Baidu Qianxi data set covers 120,142 pairs of prefecture-level cities per day for 364 such cities. For each prefecture-level city, Baidu Migration provides the following two sets of information: (1) the top 100 origination cities for the population moving to the target city and the corresponding percentages of the inflow population that originated from each of the top 100 origination cities; (2) the top 100 destination cities for the population moving out of the city and the corresponding percentages of the outflow population that go into each of the top destination cities (54). The mobility data used in this research cover the periods from January 1, 2020, to March 27, 2021 and from September 2, 2021, to April 21, 2022.

To achieve our research objectives, we converted the raw mobility data into two daily indices at the prefecture level: inflow mobility and outflow mobility. To compute the inflow mobility index for a given prefecture-level city, e.g. City A, we averaged the outflow values from all other cities directed toward City A, based on Baidu Qianxi data for a specific date.[12] Specifically, this average is derived from the percentages of outflow population originating from cities that include City A in their list of top 100 mobility destinations. Similarly, for the outflow mobility index, we followed the same procedure but substituted inflow values for outflow values in the Baidu Qianxi data. When City A implements the zero-COVID policy and assuming inter-city traffic among other cities remains constant, the share of population mobility associated with City A relative to the total population mobility of other cities is likely to decline due to imposed restrictions.

[12]In this context, the outflow mobility from other cities to City A is essentially considered as inflow mobility for City A.

This anticipated decrease would be reflected in the mobility indices we have devised.

2.2.4 PM2.5

The county-level weekly data on PM2.5 is derived from the Aerosol Optical Depth (AOD) data, which are from NASA's Global Modeling and Assimilation Office (GMAO) released Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2). Comparing to station-level PM 2.5 data, satellite data cover all the counties in China and are widely used in economic research.[13] The data is reported with a nested resolution of 50km $\times$ 60km at a hourly base. Firstly, the grid-level PM2.5 concentration is computed using the formula provided by (22). Next, to achieve a higher resolution, we split each grid into smaller grids of 5km x 6 km using an upsampling method.[14] Lastly, we adopt the *Raptor Join* method described in (119) to aggregate the data from the smaller grids into county-level for each hour and compute the weekly sum for each county.[15]

2.2.5 Night Lights

China's government has not released any county or prefecture-level GDP data for the years 2020 to 2022. Even if such data were available, there are concerns about the possibility of manipulation and over-reporting (96; 9). To obtain a consistent measure of local economic activity across China, we utilize visible lights emitted from the Earth's surface at night as a proxy — night lights (nighttime light) data have already been recognized to be capable of accurately capturing changes in local economic activity (79).[16]

We obtain the night lights data from the Visible Infrared Imaging Radiometer Suite (VIIRS) on a monthly basis,[17] covering the period from 2019 to September 2022. To filter

[13] see (59; 34; 117).

[14] If we do not upsample, there will be missing values for some counties that are smaller than 50km $\times$ 60km in size.

[15] To account for the daily air pollution's high volatility, we follow (77) and aggregate the PM 2.5 at the weekly level.

[16] Also see (76; 120; 78) and (49) for a comprehensive review of economic literature using night lights as proxy for economic actives.

[17] See (50). The raw data from VIIRS is at monthly basis.

out noise from sources such as aurora, fires, and other temporary lights, we employ a threshold of 0 and $1.5(\mu + 3\sigma)$, following (93; 60).[18] The spatial resolution of VIIRS image data is 413m, the absolute radiation values in the unit of $Watts/cm^2/sr$ (32). We use the same *Raptor Join* method describe in the PM 2.5 section to aggregate the grids at county level by month.

2.2.6 Weather Data

We obtain the weather data including precipitation and temperature from Global Historical Climatology Network form the National Oceanic and Atmospheric Adminictration (NOAA).[19] We use the inverse distance weights to calculate the daily prefecture-level weather data.

2.2.7 Daily Confirmed COVID-19 Cases

We gather the daily confirmed COVID-19 cases provided by the *Dingxiangyuan* website, which compiles official daily COVID-19 cases at the prefecture level.

2.3 Identification

Our empirical analysis relies on two sets of difference-in-differences (DiD) models to identify the impact of the zero-COVID policy on the pandemic's dynamics during local outbreaks and its subsequent influence on various measures of economic activity, including traffic mobility, air pollution, and night lights. We employ a DiD specification as our baseline regression to estimate the relative change in the outcome variable between the treated and control groups. The model is specified as follows:

$$Y_{it} = \beta D_{it} + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \varepsilon_{it}$$

[18] See Figure A3, an example of filtered data of Night Lights in March 2022 obtained from VIIRS, combine with the shapefile of China's county boundary.

[19] See (100)

where Y_{it} represents the outcome variable of interest in region (prefecture or county) i during period (day, week or month) t. D_{it} is a dummy variable indicating the treatment status in region i at time t, where it equals 1 if any community within this region is classified as a *Risk* area and 0 otherwise. Regions with *Risk* areas would be subject to the enforcement of zero-COVID policies. $\mathbf{X}_{it}$ are the control variables. μ_i represents prefecture (county) fixed effects, which control for time-invariant prefecture (county)-level factors, and θ_t represents time fixed effects, which control for shocks that are common to all regions during a given time period.

The underlying assumption for the DiD estimator is that the zero-COVID policy implementation is not driven by unobserved factors that could also systematically influence the differences in outcome variable between regions with *Risk* areas and regions without *Risk* areas. This assumption is unverifiable as it requires knowledge of the counterfactual scenario, but we can investigate whether the parallel trends assumption is satisfied before the date when any areas within these regions were classified as Risk areas. To do so, we performed an event study approach to estimate the dynamic effect of the treatment. Moreover, we can understand how long the treatment effect persists. Our model is as follows:

$$Y_{it} = \sum_{k \neq -1} \beta_k D_{it}^k + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \varepsilon_{it}$$

where D_{it}^k represents the indicator for i's treatment status at k periods relative to period t. It takes a value of 1 if region i has any areas classified as *Risk* was k periods relative to period t and 0 otherwise. We exclude $k = -1$ so that the dynamic effect is compared to the period immediately before initial treatment. The parameter of interest β_k estimates the effect of zero-COVID policy k periods after/before the implementation. We expect the pre-trends to be parallel, as β_k would not be significantly different from zero for $k \leq -2$. Intuitively, economic activities were restricted by the zero-COVID policy in the enforced regions and slowly recovered after the implementation was over, thus we expect β_k to be negative for $k \geq 0$ and converge to zero as k increases.

To investigate the heterogeneity of the effect of the Lockdown and zero-COVID policy over time, we perform separate DiD regressions and event studies for the years 2020, 2021, and 2022. As in some regions the zero-COVID policies were triggered multiple times across 2021 and 2022, we exclude the regions that have already been classified as *Risk* during 2021 from our subsample used in the analysis for year 2022.[20] As the risk level data is unavailable for 2020, we use the lockdown data from (77) to generate the treatment status for year 2020. In the following sections, we present our empirical results for different outcome variables and provide more details on the regression specifications used for our analysis.

2.4 Results

2.4.1 COVID-19 Cases

Before we examine the economic consequences of zero-COVID policies, we apply an event study approach to examine the dynamic effects of the risk level on COVID-19 cases in China, with the goal of examining the trends in COVID-19 cases before and after the implementation of the zero-COVID policy and estimating the average time it took from the launch of zero-COVID policy to when the outbreak was under control. To achieve this, we estimate the following model:

$$Case_{it} = \sum_{k=-30}^{-2} \beta_k D_{it}^k + \sum_{k=0}^{50} \beta_k D_{it}^k + \mu_i + \theta_t + \varepsilon_{it}$$

where $Case_{it}$ represents confirmed COVID-19 cases in prefecture i at date t. D_{it}^k represents the indicator for prefecture i's treatment status at k periods relative to date t.

[20] We did not exclude regions that have experienced lockdown in 2020 in any of these regressions, because, in fact, almost all prefectures in China implemented some level of restriction in mobility during the initial outbreak of the pandemic. On the other hand, the share of regions that were at *Risk* during 2021 is relatively small so the subsample after excluding these regions could still be representative.

Given the potential reverse causality between COVID-19 cases and risk level status and potential anticipation[21], we are not estimating a causal impact, but examining the correlation. The coefficient of interest β_k estimates the correlation between the status of *Risk* or non-risk k periods after/before the risk level classification and the daily confirmed COVID-19 cases. The dynamic effect results are displayed in Figure 3.

We begin by presenting the dynamic effect of the 2020 lockdown implementation in Figure 3a. Prior to the lockdown, the dynamic effect is negative. Subsequently, the effect remains positive for approximately 50 days after the initial lockdown, and reaches its peak at 40 around 21 days later, before starting to decline towards 0. It is unsurprising to observe a surge in daily confirmed cases following a lockdown, as extensive COVID-19 testing is likely to start after the lockdown is imposed when the virus has already spread for some time. As a result, the daily confirmed cases during the weeks following the lockdown tend to be higher on average than before it. Additionally, the extensive variation in the estimated dynamic effect and the predominantly insignificant estimators suggest that some prefectures prefectures may have implemented precautionary policies before potential increases in cases.

In Figure 3b and 3c, we present the results of our event study analysis for the years 2021 and 2022, respectively. Our findings suggest that the dynamic effect of zero-COVID policy on COVID-19 cases differs over the two years. Specifically, in 2021, the dynamic effect increases from day 0 to day 7 and then gradually declines, becoming negligible after day 21. In contrast, in 2022, the dynamic effect remains high for a more extended period, it takes around 25 days to control the size of the pandemic to about 5 cases, and around 50 days to decrease the magnitude close to 0. The peak of the curve is also much higher than in 2021, with an average of more than 10. Additionally, the variation of the dynamic effect in 2022 is much larger than in 2021. These findings suggest that while some prefectures were able to reduce COVID-19 cases quickly by implementing stringent measures immediately after they were classified as *Risk* regions, others found it more challenging to

[21] See (65) for a review of challenges of causality identification in COVID-19 research.

contain the spread of the virus effectively in 2022.

Overall, these findings suggest that the risk level policy in China has been effective in controlling the spread of COVID-19 in 2021, with the number of cases peaking shortly after the initial intervention and declining afterwards. However, in 2022, the emergence of new virus variants, such as the Omicron, poses challenges to the effectiveness of the policy, as it took much longer to control the pandemic in 2022 compared to 2021. Additionally, these results highlight the considerable variation in the implementation of the zero-COVID policy across different regions in China, with some prefectures experiencing a rapid decline in cases immediately after being classified as *Risk* regions, while others had a slower decline or even an increase in cases before a decline.

2.4.2 Traffic Mobility

Next, we investigate the effect of the zero-COVID policy on mobility. Our models are as follow:

$$Mobility_{it} = \beta D_{it} + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \varepsilon_{it}$$

$$Mobility_{it} = \sum_{k=-30}^{-2} \beta_k D_{it}^k + \sum_{k=0}^{50} \beta_k D_{it}^k + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \varepsilon_{it}$$

where the dependent variable $Mobility_{it}$ has two measures: inflow and outflow traffic mobility index at prefecture i on date t, taking the natural log. For the sample period of 2020, D_{it} is an indicator variable for lockdown or not.[22] For the sample period of 2021 or 2022, D_{it} is a binary variable equal to 1 if any community within this prefecture i at date t is classified as a *Risk* area and 0 otherwise. We control prefecture fixed effects by μ_i and date fixed effects by θ_t. It should be noted that the timing of the risk level classification and the adoption of corresponding zero-COVID policies may be correlated with the severity of COVID-19. We therefore include daily confirmed COVID-19 cases in the matrix of prefecture-day level controls $\mathbf{X}_{it}$. We also include weather factors in $\mathbf{X}_{it}$. The

[22] For the sample period of 2020, we use similar setting with (77)

standard errors are clustered at the prefecture level. We estimate the effect of the zero-COVID policy on mobility separately for year 2020, 2021, and 2022.

The DiD regression results in Table 2 show that the impacts of the zero-COVID policy on inflow and outflow mobility in 2021 and 2022 are significantly negative. However, the impact of lockdown on mobility in 2020 is negligible. In columns (3) and (4), the coefficients for both inflow and outflow traffic mobility during 2021 and 2022 are approximately -0.3, indicating a 30% decrease in traffic flow between a prefecture and other prefectures after it is listed as *Risk* region. This result is significant at the 1% level. In columns (1) and (2), the magnitude of the coefficient is only around -0.02, suggesting only a 2% change in traffic mobility, which is not significant. The R-squared for all regression specifications indicate that the models explain a considerable proportion of the variance, lending credibility to our estimation.

We present the dynamic effects of the lockdown and zero-COVID policy implementation on inflow and outflow traffic mobility in Figure 4 and 5, respectively. The patterns are similar for the two sets of figures within the same year. Figures 4a and 5a display the dynamic effect of lockdown on inflow and outflow mobility in 2020. There is no significant trend in the pre-treatment periods, indicating that the treatment does not affect mobility before the launch of the lockdown. Both mobility measures experienced a significantly negative effect immediately after the lockdown and stopped the decreasing trend within one week. There are sharp increases in mobility that happened during the third week after the lockdown, which may be due to the fact that the lockdown duration in 2020 was clustered around 20 days, and the mobility increase reflected the lifting of restrictions. This pattern help us to explain the insignificant lockdown effect in Table 2, On average, a significant positive rebound in traffic flow during the third week offsets the negative effects observed in the first two weeks.

In Figure 4b and 5b, we present the effect of zero-COVID on inflow and outflow mobility in 2021. The figures show a significantly negative effect that occurs immediately after the prefectures were classified as a *Risk* region, remains at a large effect size for around 15 days, and gradually returns to null around 30 days after the initial treatment. Regarding the impact of zero-COVID policy on mobility in 2022, as displayed in Figure 4c and 5c, we observe almost an identical pattern as in 2021, while the magnitude of the dynamic effects in 2022 was larger than in 2021 at its peak.

There are two possible reasons to explain this phenomenon. Firstly, it could be due to the more stringent implementation of the zero-COVID policy, which led to greater restrictions on mobility. Secondly, the release of the Travel Codes Tracker system could have also contributed to this effect by limiting travel and mobility across regions. In early 2020, despite the virus being more lethal, only individuals traveling from Wuhan were required to undergo quarantine [23]. However, in 2021 and 2022, anyone with a travel history to medium- or high-risk areas within 14 to 21 days were required to undergo mandatory quarantine at their own expense. Individuals would be tracked by the combination of Travel Code and the risk level system [24]. With the higher expected cost for traveling, it is reasonable to observe larger negative effect on the inter-prefecture traffic flow in 2021 and 2022, as compared to 2020.

In all event studies in 2021 and 2022, we observe that the pre-trend has a dip around 3 days before the enforcement of the zero-COVID policy. This suggests that people observed the COVID-19 cases and voluntarily avoided entering and leaving the region. Nevertheless, we believe that this will not harm the credibility of our DiD estimation as the scale of the pre-treatment change due to anticipation is relatively small compared to the post-treatment changes in inter-prefecture traffic mobility.

It is important to note that the impact of zero-COVID policy on traffic mobility may vary across regions, depending on the severity of the pandemic and the specific measures taken to restrict mobility. Nonetheless, our results suggest that the zero-COVID policy has been effective in restricting inter-prefecture mobility, which could contribute to controlling the

[23] See Prevention Guidance for Novel Coronavirus Pneumonia (version 4): http://www.nhc.gov.cn/xcs/zhengcwj/202002/573340613ab243b3a7f61df260551dd4/files/c791e5a7ea5149f680fdcb34dac0f54e.pdf

[24] See the reports on China's truck drivers stuck in the quarantine rules and QR trackers:https://www.reuters.com/world/china/china-truckers-use-fake-travel-records-clean-drivers-dodge-covid-rules-2022-03-30/

spread of the virus, while also negatively impacting the transportation industry and other related sectors. It should be emphasized that the measured effect is a combination of the traffic restriction effect and the "voluntary" precaution effect of the Travel Code tracker system. Furthermore, since the outcome variables are inter-prefecture traffic flows, the effect could not be attributed to within-prefecture traffic.

2.4.3 Pollution

We proceed by examining the influence associated with the zero-COVID policy on PM2.5 concentration levels in China from 2020 to 2022. Specifically, we fitted the following equations:

$$Pollution_{it} = \beta D_{it} + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \pi_{it,jm} + \varepsilon_{it}$$

$$Pollution_{it} = \sum_{k=-5}^{-2} \beta_k D_{it}^k + \sum_{k=0}^{5} \beta_k D_{it}^k + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \pi_{it,jm} + \varepsilon_{it}$$

where $Pollution_{it}$ represents the average PM2.5 concentration level at county i during week t, taking natural log. Here, we aggregate the hourly PM2.5 data into week level to average out the high volatility of the daily air pollution, following (77). For the sample period of 2020, D_{it} is a indicator for lockdown launched in county i during week t or not. For the sample period of 2021 or 2022, D_{it} is a binary variable equals 1 if any community within county i during week t is classified as a *Risk* area and 0 otherwise. We control county fixed effects μ_i and week fixed effects θ_t. $\mathbf{X}_{it}$ include daily confirmed COVID-19 cases and weather factors such as temperature and precipitation. $\pi_{it,jm}$ denotes prefecture by month fixed effects, taking value of 1 for any county i within prefecture j during month m including week t and 0 otherwise. We control prefecture by month fixed effects to account for time-variant regional conditions shared by counties within the same prefecture in a given month. The standard errors are clustered at the county level.

Table 3 reports our DiD regression results. In column (1), we replicate the estimations used in (77) and estimate the impact of lockdown on PM2.5 pollution levels in 2020. Our result is similar to theirs. In columns (2) - (5), we estimate the influence of implementing the zero-COVID policy on PM2.5 pollution levels in 2021 and 2022 and our results show an ambiguous policy effect.

Different from the lockdown effects found in column (1) of Table 3 in 2020, our findings suggest that the zero-COVID policy may not significantly reduce pollution levels in 2021. Column (2) shows a significantly positive correlation between the implementation of zero-COVID policy and PM2.5 concentration in the baseline regression. We further control for prefecture by month fixed effects in column (3), and the coefficient remains significantly positive but with a smaller magnitude. This suggests that potential time-variant prefecture-level factors that are positively correlated with the risk level status may contribute to the positive change in PM2.5 pollution level. Moreover, some time varying county-level factors might be correlated with both the probability of being classified as *Risk* region and pollution concentrations. For example, Urban counties may have a higher chance of being classified as *Risk* regions and may also experience faster increases in PM2.5 pollution levels than their rural counterparts due to their larger number of manufacturers and motor vehicles that elevate PM2.5 pollution. Overall, in 2021, county-specific growth trend of pollution appears to outweigh the influence of the zero-COVID policy.

In columns (4) and (5), we find that the policy effects become significantly negative in 2022. The zero-COVID policy reduces the PM2.5 concentration by 1.2% to 3.5%. This is expected because the zero-COVID policy imposes more stringent restrictions on economic activities in 2022. As a result, similar to the scenario in 2020, counties with *Risk* areas experienced a significant reduction in PM2.5.

To further explore the influence of zero-COVID policy on pollution levels, we present our event study analysis in Figure 6. We first replicate the model of (77) in Figure 6a for the dynamic effect of lockdown policy on pollution levels in 2020. Then we perform event studies for 2021 and 2022. Figure 6b illustrates the dynamic effect of zero-COVID on PM2.5 concentration in 2021, showing a slightly decreasing trend after the counties were classified as *Risk* areas, but with an increasing trend starting from week 3, and a positive and significant effect in weeks 4 and 5. In contrast, Figure 6c shows that in 2022, the treatment effects is negative in the first three weeks after the counties are categorized as *Risk* region. In both figures, the pre-trends are consistent with the parallel trends assumption as the coefficients prior to the treatment are all close to zero and statistically insignificant. Combining the results from our baseline DiD regression, we find that the zero-COVID policy in 2021, unlike the strict lockdown implementation in 2020, does not bring substantial improvement to air pollution levels as the restriction imposed by zero-COVID policy is limited within a county rather than the entire prefecture. However, the change in PM2.5 concentration becomes larger and more significant when counties are categorized as *Risk* regions with more stringent zero-COVID policy, as seen in 2022.

As discussed in (122), the event study approach requires relatively strong assumptions on the homogeneity of treatment effect, especially over time and across individuals, to deliver consistent estimates. These assumptions are likely to be violated in our context of zero-COVID policy, as the treatments are implemented across multiple time periods and local government could endogenously choose the stringency of their policy implementation and result in heterogeneous treatment effects. In order to overcome this potential identification issue and allow for heterogeneity in treatment effects, we apply the method proposed by (122) and present the robust estimators in our figures. In Figure 6, it can be observed that the robust estimators follow a similar pattern to the regular dynamic effect estimators and our results are robust to the potential heterogeneous treatment effects.

In conclusion, our findings reveal ambiguous effects of the zero-COVID policy on PM2.5 concentration level in 2021 and 2022. In 2021, when the zero-COVID policy was less stringent, the county-specific growth trend of pollution appears to outweigh the influence of the zero-COVID policy. In 2022, with the more stringent implementation of the zero-COVID policy, it took an average of three weeks for PM2.5 concentration to return to its original level. This suggests a corresponding three-week decrease in industrial production and traffic flow within the county. It is worth noting that during 2021 and 2022, COVID-19 containment was prioritized over environmental protection. As a result, when counties were categorized as *Risk* regions, local governments may have relaxed environmental restrictions, leading to increased pollution. This could potentially lead us to overestimate the change in pollution level associated with the implementation of the zero-COVID policy.

2.4.4 Night Lights

Finally, we present empirical evidences related to night lights (nighttime light). We use the following models:

$$NightLight_{it} = \beta D_{it} + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \pi_{it,j} + \varepsilon_{it}$$

$$NightLight_{it} = \sum_{k=-5}^{-2} \beta_k D_{it}^k + \sum_{k=0}^{5} \beta_k D_{it}^k + \mathbf{X}_{it} \times \alpha + \mu_i + \theta_t + \pi_{it,j} + \varepsilon_{it}$$

where $NightLight_{it}$ represents the night lights level at county i during month t, taking natural log. For the sample period of 2020, D_{it} is an indicator for lockdown launched in county i during month t or not. For the sample period of 2021 or 2022, D_{it} is a binary variable equal to 1 if any community within county i during month t is classified as a *Risk* area and 0 otherwise. We control county fixed effects μ_i and month fixed effects θ_t. $\mathbf{X}_{it}$ include daily confirmed COVID-19 cases and weather factors. We also include prefecture by month fixed effects for robustness, where $\pi_{it,j}$ denotes prefecture by month fixed effects, taking a value of 1 for any county i within prefecture j during month t and 0 otherwise. Similar to the effects on PM2.5, we found divergent effects of the zero-COVID policy on night lights over the sample periods, which are presented in Table 4. In column (1), we find the lockdown implementation has a significantly negative coefficient at -0.0391, which implies counties that underwent lockdowns in 2020 had a 4% decrease in night lights compared to counties that did not implement lockdowns. In columns (2) and (4), we report the estimations for the zero-COVID policy in 2021 and 2022. In 2021, we observed a positive correlation between the implementation of the zero-COVID policy and the changes in night lights, while in 2022, the zero-COVID policy reduced 14% economic activities proxied by night lights. In 2021, compared to the county-specific economic growth trend, the change in night lights associated with the less stringent zero-COVID policy in 2021 was negligible, as it only imposed restrictions in a limited number of communities within the county. However, in 2022, with the emergence of the highly contagious Omicron variant, the zero-COVID policy became stricter and seriously disrupted economic activities. We show the robustness of our results in columns (3) and (5) by controlling for prefecture by month fixed effects, and find that the coefficients remain statistically significant at the 1% level.

The dynamic effect results in Figure 7 provide further support for our argument. To allow for heterogeneity in the treatment effect over time and across treated units, we include the robust estimators of (122) in the figures. In 2020, lockdowns occurred mostly during the early phase of the pandemic, severely affecting the economic environment and consumer confidence. As shown in Figure 7a, the negative impact of lockdown on the night lights was significant and persistent, lasting more than five months after the event date, with no signs of recovery. In 2021, shown in Figure 7b, a slight increasing trend of night lights is associated with the probability that a county is categorized as a *Risk* region. A possible explanation is that counties that had more active economic performance were more likely to be classified as *Risk* regions while also experiencing faster economic recovery. However, in 2022, as shown in Figure 7c, the decreasing trend was evident, with all treatment effects negative and significant after the implementation of the zero-COVID policy. The magnitude of the negative impact continued to expand until four months after the county was categorized as a *Risk* region, with no complete recovery observed. This implies that the zero-COVID policy in 2022 brought persistent negative impacts to local economies, which may have contributed to the end of the era of zero-COVID policy and the reopening in December 2022.

It should be noted that, in Figure 7b, we observe positive pre-trend and post-trend that are significantly away from zero. This indicates that, compared to the difference in night lights between the treated and control groups in the baseline period $t = -1$, these differences in night lights are larger between two groups in periods at least two months before or after the implementation of the zero-COVID policy. As the zero-COVID policy is unlikely to bring more economic opportunities to the region due to its nature of suppressing human activities, this result could be explained by the positive correlation between the likelihood that a county will be classified as a *Risk* region and its county-specific economic growth trend in 2021. As shown in Figure A2a, only a small proportion of regions in China experienced the zero-COVID policy in 2021. It is plausible that a county in a more economically developed prefecture could enjoy a stronger recovery from the pandemic shock in 2020 and display a higher economic growth rate in 2021. Meanwhile, such prefectures were more likely to experience a pandemic outbreak in 2021. As shown in previous Section 4.1 as well as in Figure 2, the COVID-19 outbreaks in 2021 were usually on small scales and the zero-COVID policy in 2021 lasted for relatively short periods. Therefore, the persistent impact of the zero-COVID policy could be very limited in 2021. As a result, these counties could pick up the economic growth trend from the disruption of the zero-COVID policy quickly and continue with strong economic performance even after the zero-COVID policy. This potentially explains the positive estimated influence of the zero-COVID policy on night lights in 2021, as shown in columns (4)-(5) of Table 4, as well as the upward trend of dynamic effects after the treatment of zero-COVID in Figure 7b.

We provide back-of-the-envelope calculations of the GDP loss caused by the zero-COVID policy in 2022. We replicate the original dataset used in (96) and calculate the elasticity between GDP and night light. The calculation shows that a 1% change in night lights corresponds to a 0.858% change in China's GDP. Then, as shown in Panel B of Table 1, by the end of December 2022, zero-COVID policies had been implemented in 1700 out of

2853 counties in China. Finally, based on our calculation, the economic loss can be estimated as $0.077 * 0.858 * 1700/2853 = 0.039$,[25] suggesting that the zero-COVID policy resulted in a reduction in China's GDP of approximately 3.9%. Interpreting the results from this back-of-the-envelope calculation should be approached with caution due to two major limitations: (1) the estimated policy effects derived from the DiD setting may be subject to bias due to spillover effects; (2) the elasticity estimated from the data of (96) does not consider the regional heterogeneity within China. In the presence of heterogeneity between counties affected by the zero-COVID policy and counties not affected, this calculation could be inaccurate.

2.4.5 Spillover Effect Results

When two adjacent regions exhibit a close economic linkage, the implementation of human activity restrictions, such as a zero-COVID policy, in one region could exert an impact on activities in the other. This spatial correlation poses a potential bias in our difference-in-difference estimation. To isolate the spillover effects of zero-COVID policy in neighboring regions, we introduce a control variable for adjacent treated areas, following the spirit of literature on spillover effects in difference-in-difference settings (42), as well as on peer effects (62). Specifically, we define a control variable termed "Neighbors Risk" as follows:

$$Neighbors\ Risk_{it} = \frac{\sum_{j \in I \setminus i} D_{jt} R_{ij}}{\sum_{j \in I \setminus i} R_{ij}}$$

where for any two regions $i, j \in I$ at any period t, D_{jt} is a dummy variable for whether j is under the zero-COVID policy at t, R_{ij} is a dummy variable for whether a pair of prefectures i, j is neighboring. Consequently, $Neighbors\ Risk_{it}$ calculates the proportion of neighboring regions implementing a zero-COVID policy relative to all adjacent regions for a given region i at time t.

[25] We choose policy effect as .077, from Column 5 of Table 4

We incorporate this constructed "Neighbors Risk" variable, along with its lagged terms, into the primary regression models presented in prior sections. Note that to fully capture the potential spatial correlation, a spatial econometric model, such as Spatial Durbin model (SDM) is desired. Our approach only accounts for the proximate (lagged) spillover effects from zero-COVID policies in neighboring regions during 2021 and 2022. The resulting regression results for mobility, pollution, and night lights are presented in Table 5, Table 6, and Table 7, respectively.

In columns (1), (3), (5), and (7) of Table 5, we present the baseline estimates initially showcased in Table 2. Correspondingly, columns (2), (4), (6), and (8) offer estimates of local policy effects on traffic mobility that are robust to spillover influences. Across all these specifications, the local estimates exhibit only negligible variations when compared to the original findings. There is no statistically significant negative spillover effect from adjacent zero-COVID policies in 2021. However, a notable negative impact emerges in 2022, consistent with our earlier results that the stringent zero-COVID measures in 2022 exerted a more pronounced adverse effect on economic activities than those in 2021.

In columns (1) and (3) of Table 6, we offer the baseline estimates for pollution outcomes from Table 3, while columns (2) and (4) include regression results accounting for spillover effects. No substantive changes are observed compared to the original estimations, and negative, statistically significant spillover effects are found for both 2021 and 2022. Given that PM2.5 concentration data are extracted from satellite image and aggregated at the county level, it is plausible that the implementation of a zero-COVID policy in a neighboring county could reduce pollution levels in adjacent areas due to restricted traffic mobility and manufacturing.

In columns (1) and (3) of Table 7, we present the baseline estimates for night light data from Table 4 and include spillover-adjusted regression results in columns (2) and (4). Again, the estimates remain substantively unchanged compared to the original findings, and no statistically significant spillover effects on night lights are observed for either 2021 or 2022. This suggests that long-term economic activities, as reflected by night light data, are unlikely to be influenced by zero-COVID policies in nearby regions.

2.4.6 Synthetic Diff-in-Diff Results

As mentioned in Section 4.4, potential selection bias may exist within the treated sample. Specifically, regions with greater economic development could be more susceptible to experiencing COVID-19 outbreaks, thereby making it more likely for them to implement the zero-COVID policy and consequently be included in the treatment group. In estimating the impact of zero-COVID policy implementation on economic outcomes like pollution and night lights, uncontrolled county-level time-varying trends could raise concerns regarding the validity of our estimated results.

To enhance the comparability between the treatment and control groups in our empirical examination of pollution and night lights, we conduct several auxiliary regressions employing the Synthetic Difference-in-Differences (SDID) method, a fusion of the Difference-in-Differences and Synthetic Control methodologies (10). The SDID approach assigns weights to individual fixed effects and time fixed effects to approximate the pretrends between the treatment and control groups, thereby mitigating the reliance on the parallel trends assumption and generating more stable and robust estimates. It is noteworthy that, to comply with the SDID framework, we keep a balanced sample, resulting in a reduced sample size. We also disclose the outcomes of our primary regressions utilizing the balanced sample in Table A1 and Table A2 for reference.

We present our SDID estimations for pollution and night lights in Table 8 and Table 9, respectively. In column (1) of Table 8, the impact of lockdowns on PM2.5 concentration remains significantly negative. In column (2) of Table 8, the estimated changes in PM2.5 following the initiation of zero-COVID policy retain the same sign as the original estimate. In column (3), the estimated coefficient shifts from negative to positive, though with a relatively small magnitude. These outcomes align with our prior findings presented in the dynamic effect results of Figure 6. Specifically, local pollution showed a marked decline post-lockdown in 2020; its trend began to rise a few weeks following the implementation of the zero-COVID policy in 2021; and in 2022, the pollution exhibited a short-lived dip but did not sustain it.

In Table 9, we observe similar results for changes in night lights correlated with zero-COVID policy, compared with the original estimates in Table 4. The estimated coefficient for 2020 remains negative, though its statistical significance diminishes, while the results for 2021 and 2022 retain their original signs and are statistically significant. In summary, despite potential confounding factors involving the relationship between the implementation of zero-COVID policy and pre-treatment economic trends, our SDID estimations reaffirm the robustness of our primary regression outcomes and are consistent with our other findings.

2.5 Conclusions

In this paper, we provide evidence on the economic impacts of the zero-COVID policy implemented by the Chinese government as a pilot experiment in using big data for country management from 2020 to 2022. We employ a difference-in-differences specification with a novel dataset of China's COVID-19 risk level system. First, we find that the zero-COVID policy in China effectively contained COVID-19 transmission within a 21-day window in 2021. However, controlling virus transmission took twice as long with the emergence of the Omicron variant in 2022. Second, the zero-COVID policy led to a 30% reduction in inflow and outflow mobility, indicating a potential negative impact on the transportation industry and related sectors. Third, our study indicates that the zero-COVID policy had a negligible effect on pollution levels in 2021. Nevertheless, it led to a decrease in PM2.5 concentration in the estimated range of 1.17% to 3.47% in 2022. Lastly, the evidence reveals that the zero-COVID policy had trivial impacts on night lights in 2021, which was overshadowed by the strong economic performance due to the recovery effect. However, we discover a significant reduction in economic activities proxied by night lights, ranging from 7.7% to 14%, as a result of the implementation of the zero-COVID policy in 2022. We calculate that the zero-COVID policy resulted in a reduction of approximately 3.9% in GDP.

Several other countries pursued an elimination strategy like China, with strict border controls and lockdowns to keep the virus at bay, for example, New Zealand, Australia, Singapore, Vietnam, and Thailand. Studies generally show that COVID-19 has had a negative impact on these economies, especially for the countries that rely heavily on tourism and international trade (47; 24; 58; 109). Most countries experienced an economic contraction during the initial stage of the pandemic, but were able to have a quick rebound since their proactive response to the pandemic had effectively minimized cases infected. An exact comparison of economic impacts between China and these countries, however, is not feasible because the strictness of containment policies enforced by different countries varies, and some countries shift their strategies in response to changing circumstances at different times.

Overall, our findings offer important insights into the effectiveness and limitations of the zero-COVID policy in controlling the spread of COVID-19, as well as its impact on various aspects of the economy and society. These insights can inform the design and implementation of big data-driven public health policies that aim to reduce the impact of public health crises and minimize economic costs in China.

2.6 Figures and Tables

2.6.1 Figures

Figure 2.1: Daily Confirmed Cases v.s. Number of Counties with *Risk* (excluding Shanghai)

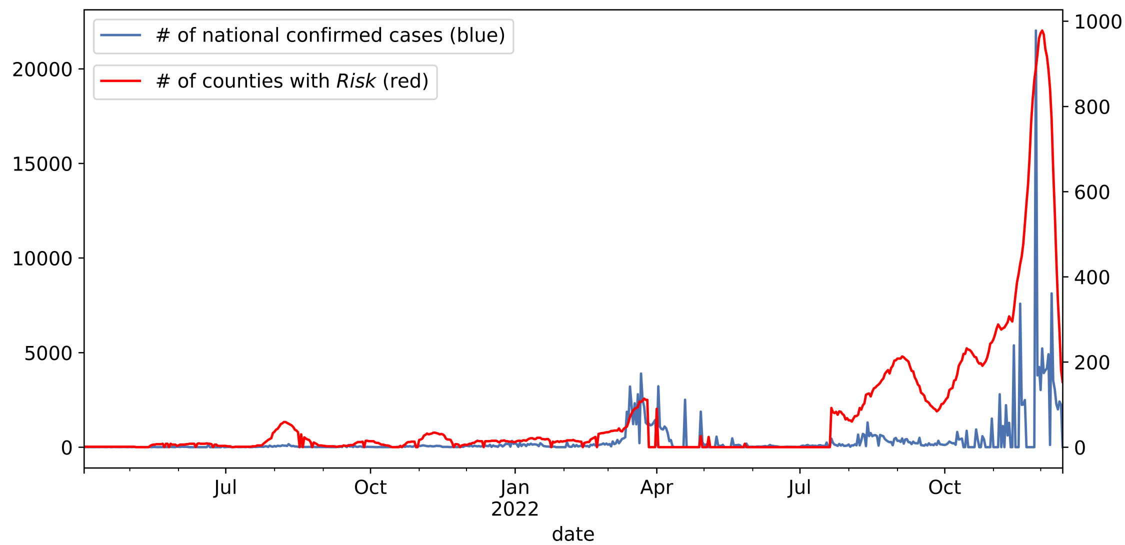


Figure 2.2: Distribution of *Risk* Duration per County

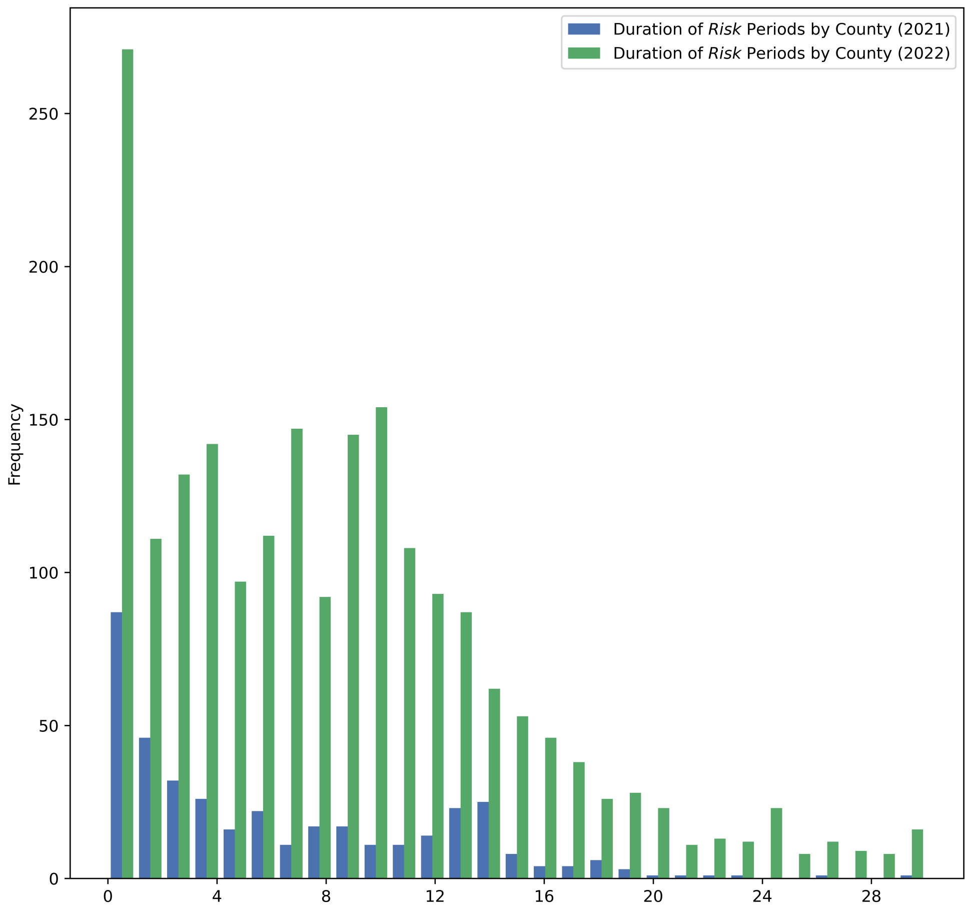
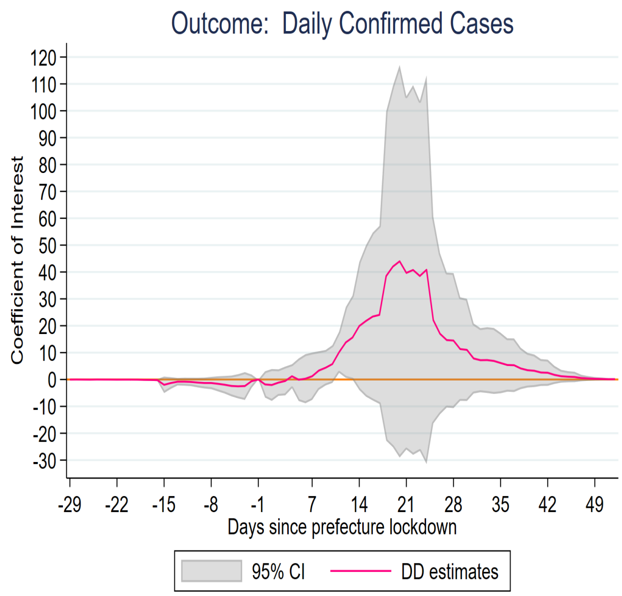
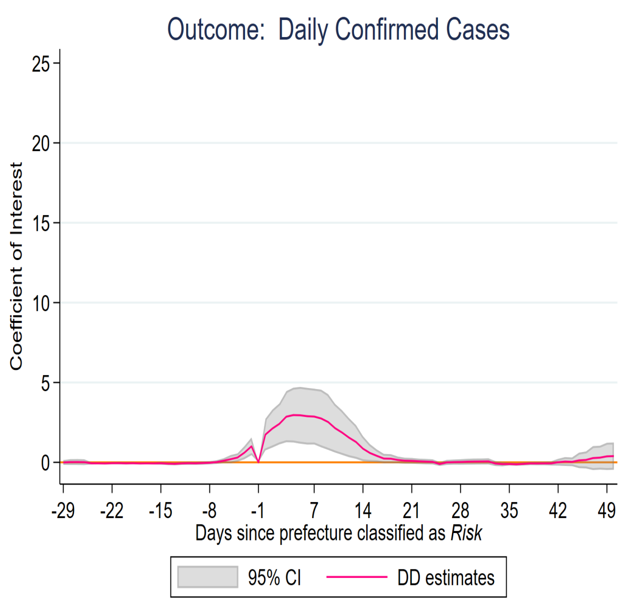


Figure 3: Event Study: Daily Confirmed Cases

(a) Lockdown 2020



(b) *Risk* 2021



(c) *Risk* 2022

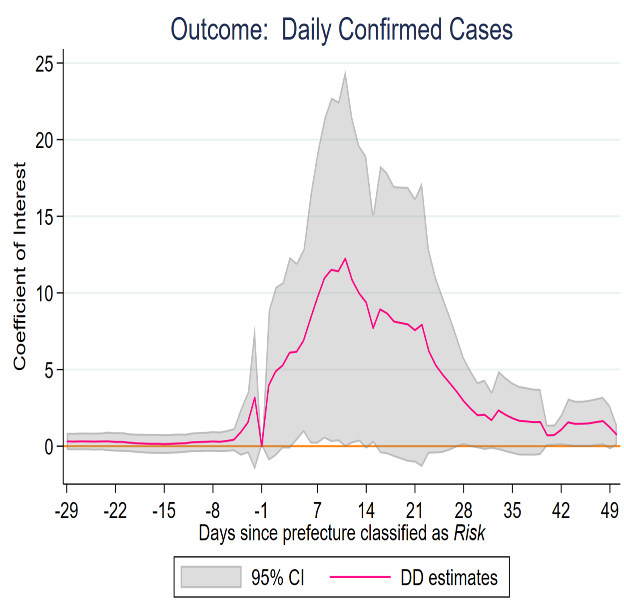
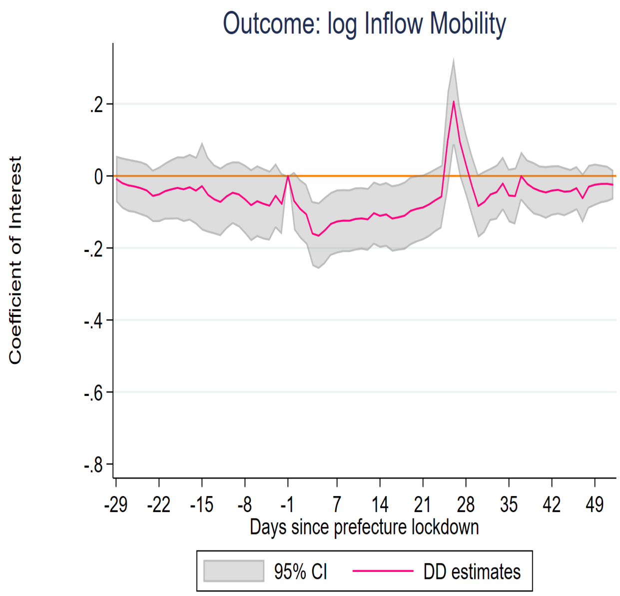
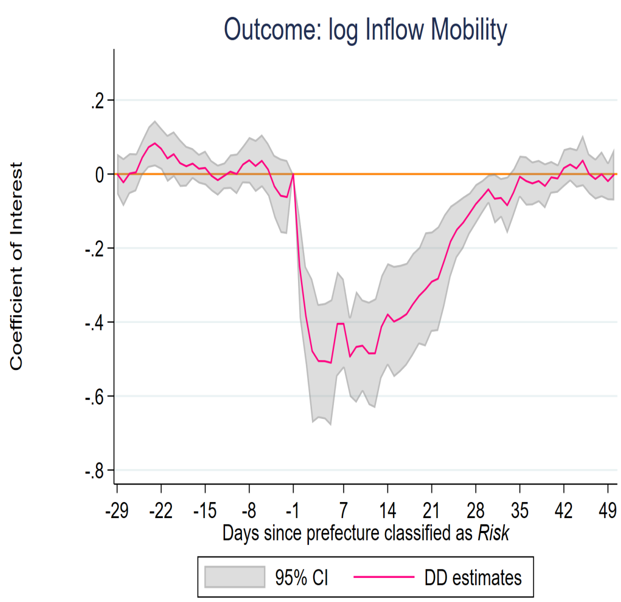


Figure 4: Event Study: Inflow Mobility

(a) Lockdown 2020



(b) *Risk* 2021



(c) *Risk* 2022

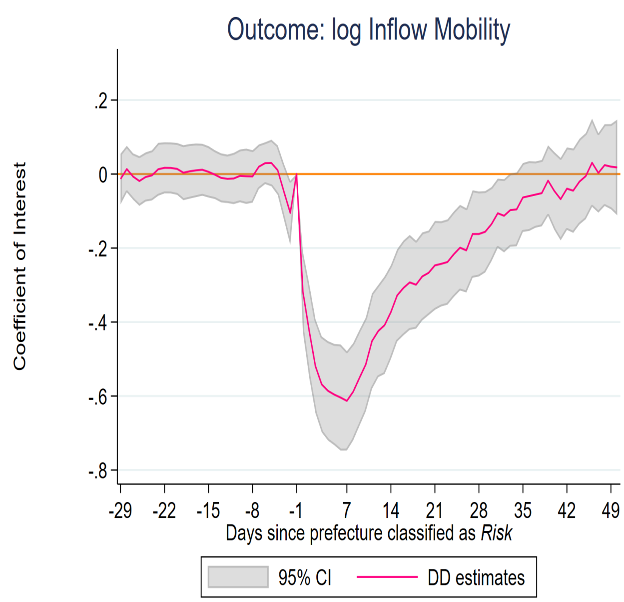
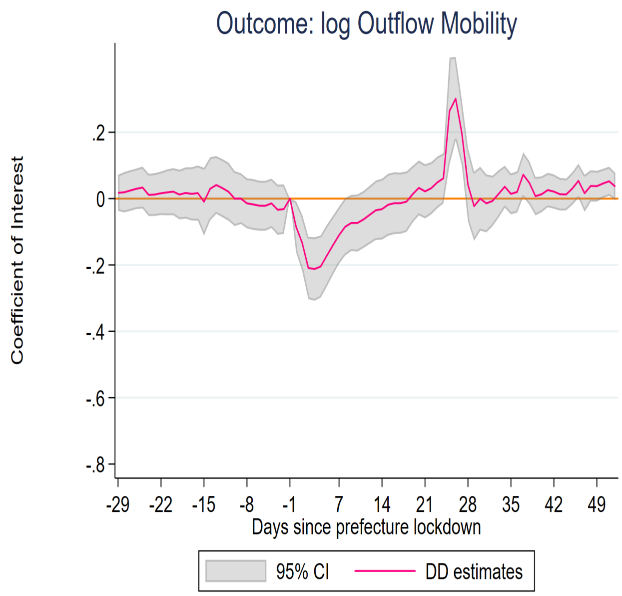
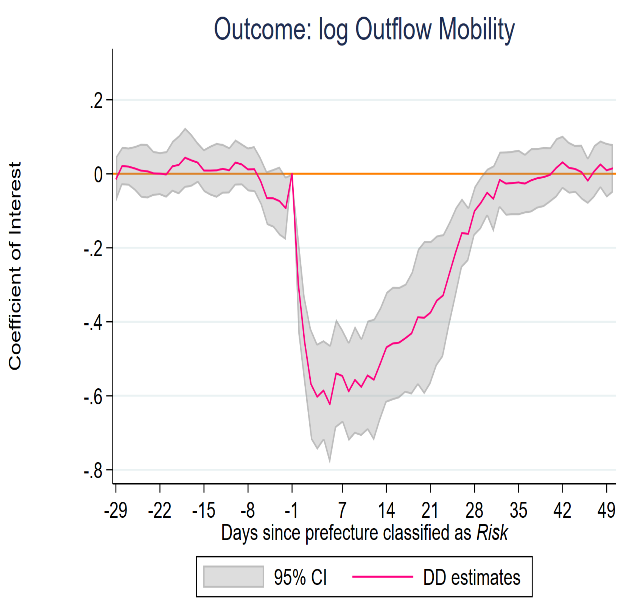


Figure 5: Event Study: Outflow Mobility

(a) Lockdown 2020



(b) *Risk* 2021



(c) *Risk* 2022

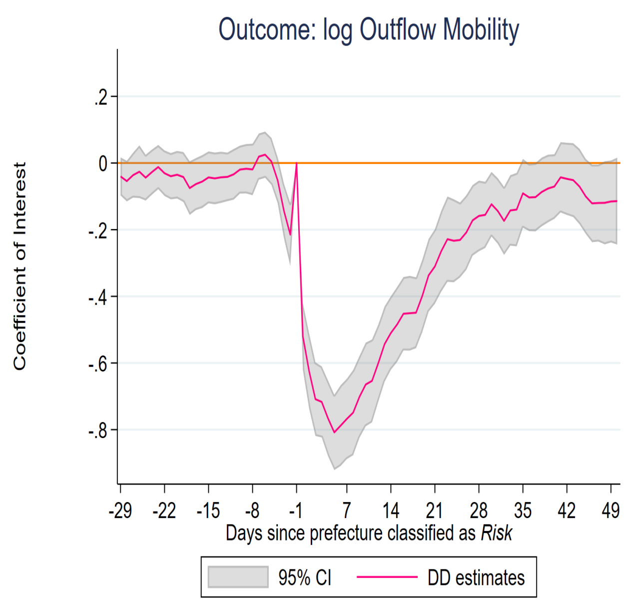
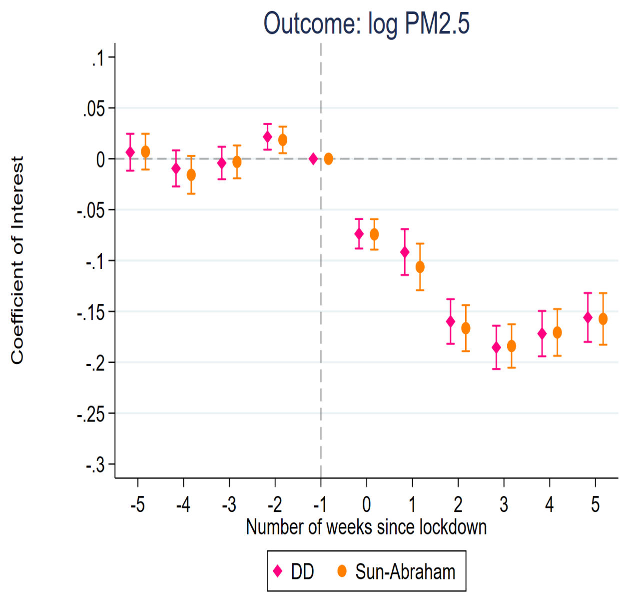
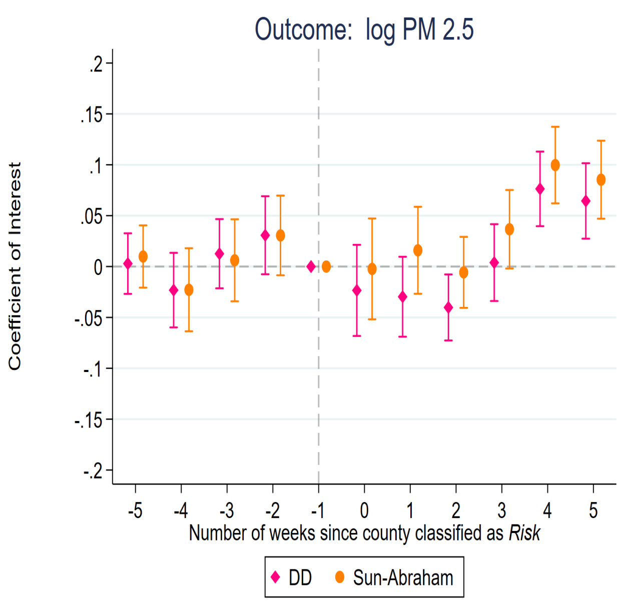


Figure 6: Event Study: PM2.5

(a) Lockdown 2020



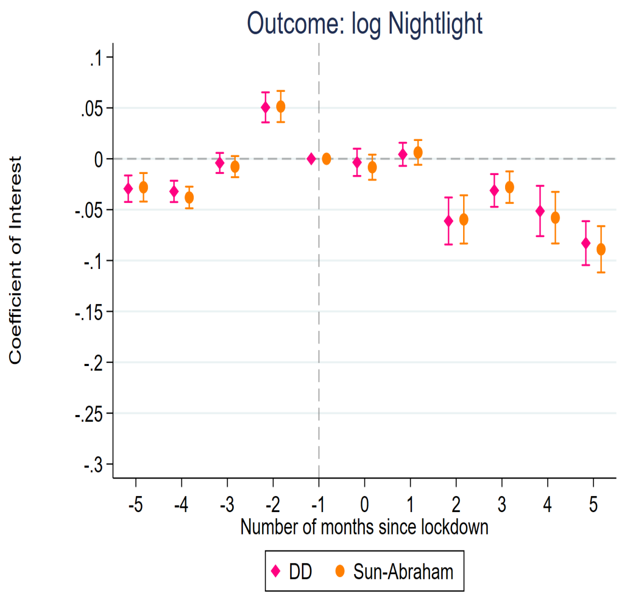
(b) *Risk* 2021



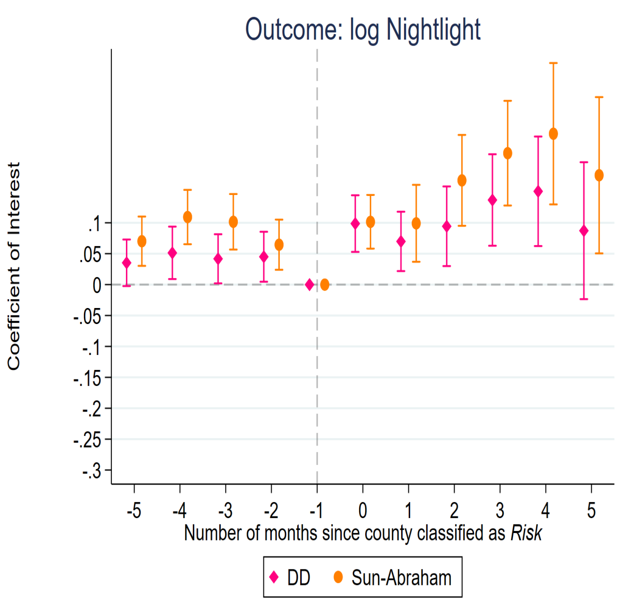
(c) *Risk* 2022

Figure 7: Event Study: Night Light

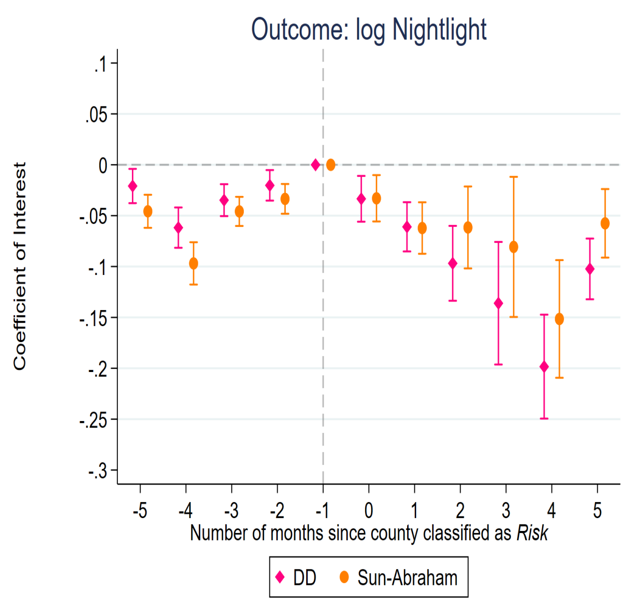
(a) Lockdown 2020



(b) *Risk* 2021



(c) *Risk* 2022



2.6.2 Tables

	Obs	Mean	Std.Dev	Min	Max
Panel A: County Panel	ref.				
Classified as *Risk* (County)	1777419	0.026	0.159	0.0	1
Night Lights (monthly) ($Watts/cm^2/sr$)	45350	2.420	4.457	0.1	53
PM2.5 (weekly) (μ/m^3)	253352	26.665	15.070	0.4	394
Panel B: County by Dec15,2022	ref.				
Cumulative Days Classified as *Risk* (County)	2853	16.095	23.231	0.0	243
Cumulative Days Classified as *Risk* (Exclude Never Treated)	1700	27.011	24.716	1.0	243
Panel C: Prefecture Panel	ref.				
Classified as *Risk* (Pref)	229264	0.074	0.262	0.0	1
Share of counties Classified as risk (Pref)	229264	0.025	0.117	0.0	1
Num of Counties Classified as *Risk* (Pref)	229264	0.200	1.025	0.0	35
Daily Confirmed COVID Cases	657218	0.545	47.862	0.0	23718
Inflow Mobility	179041	0.281	0.313	0.0	4
Outflow Mobility	175695	0.281	0.321	0.0	5
Panel D: Prefecture by Dec15,2022	ref.				
Cumulative Days Classified as *Risk* (pref)	368	46.220	41.815	0.0	250
Cumulative Confirmed COVID Cases	356	1001.298	5239.218	1.0	64978

Table 1: Statistical Summary

Table 2: Mobility Regression Results

	2020		2021		2022	
	(1)	(2)	(3)	(4)	(5)	(6)
	log Mobility Inflow	log Mobility Outflow	log Mobility Inflow	log Mobility Outflow	log Mobility Inflow	log Mobility Outflow
Lockdown	-0.0137	-0.0283				
	(0.0442)	(0.0366)				
Risk			-0.287***	-0.350***	-0.292***	-0.323***
			(0.0344)	(0.0431)	(0.0385)	(0.0383)
R-squared	0.835	0.849	0.968	0.963	0.888	0.883
Observations	23231	23234	30060	30060	38352	38352
Mean of Mobility	0.277	0.277	0.284	0.284	0.234	0.235
Controls	✓	✓	✓	✓	✓	✓
Prefecture FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Pollution Regression Results

	(1) log PM2.5 2020	(2) log PM2.5 2021	(3) log PM2.5 2021	(4) log PM2.5 2022	(5) log PM2.5 2022
Lockdown	-0.162***				
	(0.00956)				
Risk		0.0842***	0.0159***	-0.0347***	-0.0117***
		(0.0153)	(0.00576)	(0.0100)	(0.00320)
R-squared	0.870	0.773	0.873	0.749	0.882
Observations	42750	99750	99050	145944	144864
Mean of PM2.5 (Weekly Average)	31.52	25.83	25.83	26.94	26.94
Controls	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓
Prefecture × Month FE			✓		✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Night Lights Regression Results

	(1) log NightLight 2020	(2) log NightLight 2021	(3) log NightLight 2021	(4) log NightLight 2022	(5) log NightLight 2022
Lockdown	-0.0391***				
	(0.00643)				
Risk		0.0812***	0.219***	-0.139***	-0.0771***
		(0.0229)	(0.0241)	(0.0135)	(0.0129)
R-squared	0.969	0.980	0.965	0.962	0.980
Observations	40103	19598	19991	23838	23341
Mean of Nightlight (Monthly)	2.614	2.268	2.268	2.354	2.354
Controls	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓
Month FE	✓	✓		✓	
Prefecture × Month FE			✓		✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Mobility Spillover Results

	2021				2022			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	log Mobility Inflow	log Mobility Inflow	log Mobility Outflow	log Mobility Outflow	log Mobility Inflow	log Mobility Inflow	log Mobility Outflow	log Mobility Outflow
Risk	-0.287***	-0.286***	-0.350***	-0.349***	-0.292***	-0.295***	-0.322***	-0.321***
	(0.0344)	(0.0354)	(0.0431)	(0.0446)	(0.0385)	(0.0394)	(0.0384)	(0.0398)
Neighbors Risk		-0.0309		-0.0620		-0.0849*		-0.195***
		(0.0345)		(0.0500)		(0.0495)		(0.0429)
R-squared	0.968	0.969	0.963	0.963	0.888	0.885	0.883	0.880
Observations	30060	29970	30060	29970	38352	36585	38352	36585
Mean of Mobility	0.284	0.284	0.284	0.284	0.234	0.234	0.235	0.235
Neighbors Risk Lag 7 days		✓		✓		✓		✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Prefecture FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Pollution Spillover Results

	(1)	(2)	(3)	(4)
	log PM2.5 2021	log PM2.5 2021	log PM2.5 2022	log PM2.5 2022
Risk	0.0159***	0.0161***	-0.0117***	-0.00917***
	(0.00576)	(0.00624)	(0.00320)	(0.00326)
Neighbors Risk		-0.0982***		-0.0464***
		(0.0306)		(0.00868)
R-squared	0.873	0.878	0.882	0.880
Observations	99050	93060	144864	139072
Mean of PM2.5 (Weekly Average)	25.83	25.59	26.94	26.72
Neighbors Risk Lag 2 weeks		✓		✓
Controls	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Week FE	✓	✓	✓	✓
Prefecture × Month FE	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Night Lights Spillover Results

	(1) log NightLight 2021	(2) log NightLight 2021	(3) log NightLight 2022	(4) log NightLight 2022
Risk	0.0812***	0.0997***	-0.0771***	-0.0703***
	(0.0229)	(0.0247)	(0.0129)	(0.0129)
Neighbors Risk		-0.0386		0.00277
		(0.0393)		(0.0178)
R-squared	0.980	0.979	0.980	0.979
Observations	19598	17196	23341	20626
Mean of Nightlight (Monthly)	2.268	2.317	2.354	2.254
Neighbors Risk Lag 1 month		✓		✓
Controls	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Prefecture × Month FE	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Pollution SDID Results

	(1) log PM2.5 2020
Lockdown	-0.115***
	(0.00839)
Observations	42750
Mean of PM2.5 (Weekly Average)	31.52
Controls	✓
County FE	✓
Week FE	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Night Lights SDID Results

	(1) log NightLight 2020
Lockdown	-0.00401
	(0.00465)
Observations	23790
Mean of PM2.5 (Weekly Average)	2.385
Controls	✓
County FE	✓
Week FE	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.7 Appendix

2.7.1 Appendix A: China's COVID Risk Level Dataset

In order to comply with the Prevention Guidance for Novel Coronavirus Pneumonia (version 5),[26] starting from March 2020, the State Council of China began to release a national COVID risk level system on a regular basis through their website. This system categorizes communities within the 2853 counties into high, medium, or low-risk groups on a daily basis. In specific, the risk level is reported by local governments and compiled by National Health Commission of China.

This website had two access interfaces. Interface A on the left column of Figure A1 is a search engine that allows users to obtain communities' risk level results for a specific county by entering its name. Interface B, located in the right column, displays all counties that have communities classified as *Risk* along with their corresponding community names. Counties that do not appear on this list are considered non-risk areas.[27]

We started risk level data collection through interface B since April 02, 2021 and ended by Dec 15, 2022.[28] [29] The China COVID Risk Level Dataset contains daily risk level information for 2853 counties from April 02, 2021 to December 15, 2022. This dataset is the most systematic compilation of China's risk level classification during 2021 and 2022. Please visit this link to access the dataset.

[26] Prevention Guidance for Novel Coronavirus Pneumonia (version 5): http://www.nhc.gov.cn/jkj/s3577/202002/a5d6f7b8c48c451c87dba14889b30147.shtml and a follow up guidance: http://www.gov.cn/zhengce/zhengceku/2020-04/16/content_5503261.htm

[27] The web links for both pages have already expired. Interface A: bmfw.www.gov.cn/yqfxdjcx/index.html and Interface B: bmfw.www.gov.cn/yqfxdjcx/risk.html

[28] The weblink of interface B expired on Dec 15, 2022. But the weblink of interface A was still active until Dec 25, 2022, we collected the data between Dec 15 to Dec 25 through a third party website, http://bj.bendibao.com/ but did not integrate the last 10 days data into our dataset yet.

[29] We thank open-source projects *BeautifulSoup* and *Selenium*.

Figure A1: Demo of State Council's website for the Risk Level System.

(a) Interface A

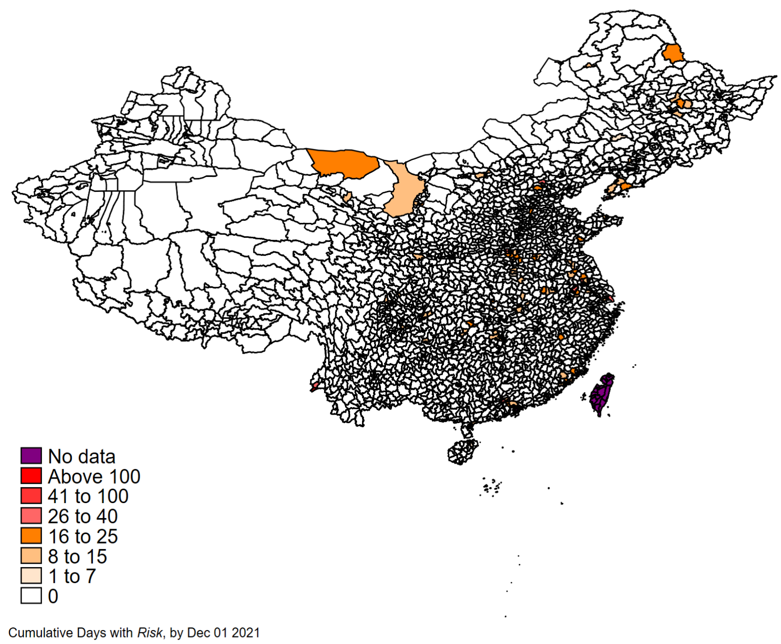
(b) Interface B

Figure A1:

Notes: The web links for both pages have already expired.
Interface A: bmfw.www.gov.cn/yqfxdjcx/index.html
and Interface B: bmfw.www.gov.cn/yqfxdjcx/risk.html

Figure A2: Geographical Distribution of counties with *Risk*

(a) *Risk* 2021



(b) *Risk* 2022

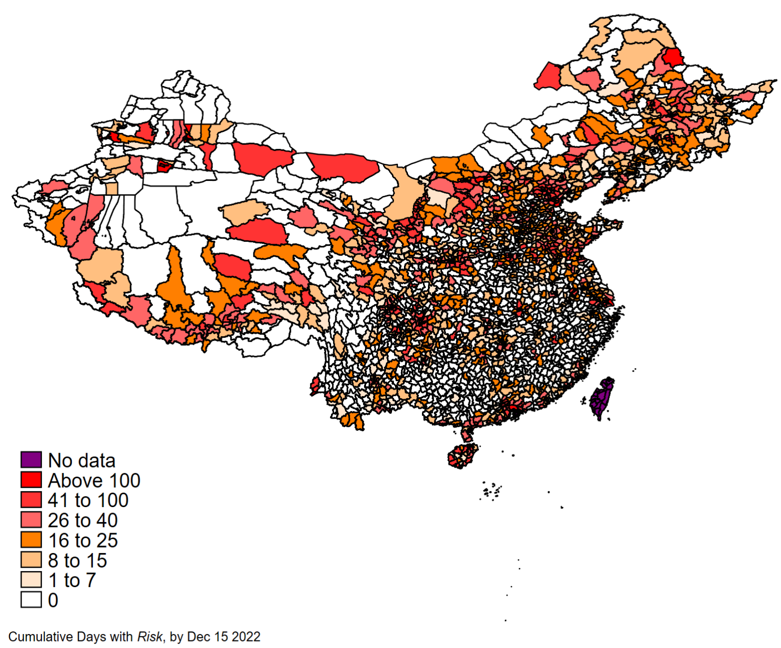


Figure A3: Night Lights in March 2022

Figure A3:

Note: This is the filtered data of Night Lights in March 2022 obtained from VIIRS, combine with the shapefile of China's county boundary.

Table A1: Pollution Balanced Sample Regression Results

	(1) log PM2.5 2020	(2) log PM2.5 2021	(3) log PM2.5 2021
Lockdown	-0.162***		
	(0.00956)		
Risk		0.0842***	-0.0344***
		(0.0153)	(0.0101)
R-squared	0.870	0.773	0.749
Observations	42750	99750	145584
Mean of PM2.5 (Weekly Average)	31.52	25.83	26.93
Controls	✓	✓	✓
County FE	✓	✓	✓
Week FE	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Night Lights Balanced Sample Regressions Results

	(1)	(2)	(3)
	log NightLight 2020	log NightLight 2021	log NightLight 2021
Lockdown	-0.0669***		
	(0.00951)		
Risk		0.0777***	-0.137***
		(0.0249)	(0.0261)
R-squared	0.966	0.981	0.953
Observations	23790	12648	11000
Mean of Nightlight (Monthly)	2.385	2.083	1.844
Controls	✓	✓	✓
County FE	✓	✓	✓
Month FE	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

Cost of Zero-Covid: Effects of Anti-contagious Policy on Labor Market Outcomes in China

3.1 Introduction

Most countries around the world have taken various containment measures to limit the spread of COVID-19, including closing public gathering places, limiting transportation services, implementing stay-at-home mandates or lockdowns, and so on. However, consensus regarding the economic impact of the anti-contagious measures has not been achieved. Some critics of anti-contagious policies claim that they slow economic growth and hurt consumer spending, while proponents argue that the economy would still deteriorate without these measures due to the fear of viruses (67) .

In this paper, we examine the effect of anti-contagious policies on labor market outcomes. One of the substantial challenges in evaluating the costs and benefits of different anti-contagious policies is to distinguish between the economic damage caused by the anti-contagious measures and the direct public health shock. In the face of this unprecedented pandemic, most countries are unable to contain the emergence of new cases right after implementing disease prevention policies, thus leading to a persistent public health shock as well as the impacts of the mitigation policies (66).

Our focus is on the anti-contagious policy in China, also known as zero-Covid policy. The setting is particularly relevant because of China's strong intention and ability to combat the pneumonia outbreak. After the outbreak of the pandemic in Wuhan, China quickly adopted the most stringent disease prevention and control policies, which effectively stopped the spread of the virus in most areas (115; 53; 90; 82; 121). This zero-Covid policy adopted by the Chinese government requires immediate disease prevention measures after finding new cases, as well as a 14-day observation window before lifting the restrictions. When new COVID-19 cases arise, this approach aims to eliminate the virus as soon as possible. Therefore, the economic fallout is mainly due to the anti-contagious policy in China, rather than the public health shock.[1]

Another challenge to accurately estimate the impact of the zero-Covid policy is the spillover effect. As soon as a prefecture implements stringent anti-contagious measures, such as city lockdowns, human mobility will fall dramatically and business will cease not only within the focal region but also between it and other regions. This implies that economic activities in a prefecture could be influenced by the zero-Covid policies of nearby regions if their economic connections were strong before the outbreak, which would bias the effect of anti-contagious policy. We control the spillover effect by controlling every prefecture's nearby zero-Covid policy duration, and our results show that the estimated local policy effect is not driven by the spillover effect.

Our paper exploits the policy design and employs a generalized Difference-in-Differences (DiD) strategy to estimate the causal effect of the duration of zero-Covid policy on labor market outcomes. The estimation result indicates that an 10% increase (in average 3.7 days) in the policy duration causes the individual unemployment probability increase by

[1]Evidences (97; 131) suggest that during 2020, China did not experience a significant increase in excess mortality, indicating that there was neither a direct health impact of the pandemic nor indirect health effects caused by the stringent anti-contagious policies or a dysfunctional medical system.

around 0.1. [2] Our estimated policy impact is the marginal treatment effect on the probability of unemployment for individuals who stayed in regions that experienced a longer zero-Covid policy duration compared to those who were located at regions with a shorter policy duration. It is noteworthy that the magnitude of the unemployment probability for an individual is incomparable to the aggregate unemployment rate as only a small proportion of the population was in the treatment group which has a larger duration of zero-Covid policy, and the aggregate unemployment rate could have a much smaller fluctuation during the pandemic. Furthermore, our result disentangles the impact of the anti-contagious policy on labor market from the public health shock and the spillover effect from nearby regions. We provide the evidence of the associative economic cost of China's zero-Covid policy for eliminating the pandemic in 2020.

This paper relates to several strands of literature. First of all, it contributes to the increasingly large empirical literature on identifying causal impact of COVID-19 policies on labor market.[3] [4] (73) apply a Difference-in-Differences structure to estimate the causal effect of social distancing policies on labor market in US during the early phase of the pandemic. Their counterfactual estimate shows that social distancing policies explain about 60% of the realized decline in employment, while without the social distancing policies it is likely to endure a more severe public health problem which could in turn deteriorate the labor outcomes. (81) use a measure of people's mobility with policy instruments and implements a 2SLS estimate on the effect of restricted mobility induced

[2]Our measurement of unemployment is similar to U-4 unemployment definition – a worker is "unemployed" if unemployed or discouraged. In Section 3.1, we will explain with details how we measure an individual employment status.

[3]A survey on this topic could be found at (21); an overview on the global labor market influence could be found at (107; 108).

[4]The pandemic causes a general negative effect on labor outcomes including employment, hours worked and income, with heterogeneous magnitudes across different countries and among different groups of workers. (43), (101), (91), (57), and (25) analyze the pandemic impact on the US labor market and household income; (139) investigates the working hour and income change in Netherlands; (6) investigates the labor market disruption in Norway; (3) documents immediate impact of the pandemic on the employment status for workers in UK, US and German; (19) investigates the shock on the US immigrant employment ; (26) and (89) analyzes the pandemic shock on the US labor market from both of supply and demand sides by using the payroll data and real-time establishment-level data; (17) investigates the idiosyncratic impact of the pandemic for different demographic groups in US; (40) investigates the heterogeneous impact on the labor market based on a granular level real-time private company data.

by policy on labor market outcomes. Their use of policy as an instrument helps create the exogenous change in mobility. (12; 13) provide a benchmark for the marginal unemployment rate change in the number of infections where there is no mandated lockdowns in South Korea.

Our choice of treatment, identification setting and unique context provide credibility to identify causal effects of COVID policies, by tackling down possible challenges.[5] First, our framework enables us to construct a conditional exogenous treatment — the duration that a prefecture with emerged COVID-19 case in the past 14 consecutive days. Under the assumption that the date of detecting a positive COVID-19 case is random, and by controlling for the prefecture fixed effect, we could causally interpret the impact of China's anti-contagious policies on labor outcomes. Second, our identification design allows us to analyze the disentangled impact of zero-Covid policy on labor market, rather than combined public health shock and derivative voluntary precautions.[6] There are no other studies, to our best knowledge, that conduct analysis on the impact of anti-contagious policies on labor market without the existence of the public health shock caused by the pandemic. Third, in the context of China, our result provides a benchmark for the marginal change in labor market outcomes where the region implements anti-contagious measures without a significant scale of the pandemic. Furthermore, the unprecedented anti-contagious policy launched by China leaves little chance for anticipation.

In addition, this paper supplements the existing research on the economic cost of the COVID-19 pandemic in China by focusing on the effectiveness of zero-Covid policy package rather than a specific policy. Recent literature on COVID-19 impact in China (136; 77; 114; 33) focuses on the influence from city lockdowns, while this paper identifies the zero-Covid policy effect for a wide spectrum of containment measures at different

[5](65) discussed several potential threats to the validity of DiD designs when used to identify the causal effects of COVID-19 policies, such as packaged policies, voluntary precautions, anticipation and spillovers.

[6]Compared to most countries analyzed in the literature, China experienced very limited pandemic surge in 2020 after the very first outbreak in Wuhan. As Chinese society is not largely influenced by the health threat of the pandemic, the zero-Covid policy contributes the most to the observed labor market disruption. Thus, our result sheds light on the isolated policy effect on the economic activity during the pandemic.

intensities, e.g. lockdowns, regional quarantines, closure of public places, transportation restrictions, etc. During the early period of China's anti-Covid campaign, many prefectures implemented lockdowns to block the spread of the virus quickly and efficiently, while more prefectures which experienced mild outbreak of the epidemic chose less stringent measures to contain this public health crisis. Our estimations capture the impact of zero-Covid policy not limited to the lockdown, but any anti-contagious measures will be counted. Our work unveils the unclear question that how much impact did these non-lockdown measures impose on the labor outcomes. Furthermore, to the best of our knowledge, our paper is the first paper to complement the literature by using the China Family Panel Studies, which is a nationally representative social survey for China.

Finally, this paper is related to the research on human mobility restriction in response to pandemic threats. Many countries implemented measures that limit the human mobility flows to stop the transmission of infectious diseases (45; 14; 132; 30). Meanwhile, the evaluation of restrictions on human mobility remains obscure for two major concerns, the negative economic impacts and the effectiveness of such policies in containing the pandemic. It is also hard to disentangle the impact of human mobility from other channels (56; 80). In this paper, we provide an estimation of the disentangled effect in the labor market of one specific mobility restriction policy, the zero-Covid strategy, which is proved to be effective in delaying and containing the spread of the pandemic (53). Our results contribute to the evaluation of human mobility restriction policy by providing a reference of the potential economic cost of halting the pandemic in perspective of labor outcomes. Our work provides a benchmark for the scenario that the pandemic was constrained effectively after the initial outbreak by fast and stringent containment measures so that the only labor market disturbance were caused by the anti-contagious policies rather than the public health shocks which most countries suffered from during the pandemic.

This paper is organized as follows. Section 2 introduces China's anti-contagious policies after the outbreak of COVID-19. Section 3 summarizes the individual survey data, COVID-19 data and regional economic data. Section 4 displays our identification strategy.

Section 5 discuss our estimation results.

3.2 Background

China's zero-Covid policy[7] consists of two components, *stringent clearance* and *dynamic clearance*. Stringent clearance includes policy responses such as quarantine, lockdown and traffic restriction. However, in regions with mild outbreaks, dynamic clearance policies with fewer restrictions on human mobility are implemented. At the initial outbreak of the pandemic, from January to February 2020, the stringent clearance prevailed in areas with COVID cases. As the government started aiming to resume work and production after Feb 17, 2020, zero-Covid policy became a hybrid between stringent clearance and dynamic clearance.

3.2.1 First Phase: Stringent Clearance

China implemented a series of unprecedented lockdowns and non-pharmacological anti-contagious policy measures in an effort to halt the spread of COVID-19 since January 23, 2020.[8] Based on Figure A1, by January 25, 30 out of China's 31 provinces had enacted First level emergency response, measures taken including case isolation, suspension of public transportation and public space closure, etc. (115; 128). Local governments reacted with *stringent clearance* policies in response to the unprecedented national emergency. The entire Hubei province implemented the lockdown in Jan 24, and its residents could not leave their prefectures. There were also strict anti-contagious policies implemented in other provinces, including a partial lockdown, a ban on traffic leaving and a 14-day

[7]Note that the term "zero-Covid" in (33) only refers to the *stringent clearance* in our paper.

[8]According to Emergency Response Law of the PRC, the emergency events are classified into 4 levels, First as extreme important and Fourth as normal. The First level emergency response is coordinated by the central government, the Second level is led by province government, the Third level is led by the prefecture government and the Fourth level is led by county government. There is no specific instruction on how to response to different emergency levels (i.e, lockdown or travel restriction), so this province level indicator is considered as a bellwether for province government's attitude towards COVID. Ironically, as shown in Figure A1, Hubei province, the center of COVID outbreak, only acted the Second level emergency response on Jan 24, and upgraded to the First level on the next day.

self-quarantine period for visitors. According to (115), up to 14,000 health checkpoints were set up at ferry and highway service centers. By February 16, more than 250 prefectures rolled out measures such as "closed management of communities", "family outdoor restrictions", "only one person of each family may go out for shopping once every 2 days", "tracing and quarantining close contacts of suspicious cases" and so on.[9] Under such stringent clearance policies, in January and February, economic activities were rigorously suppressed (53). In Appendix 7.1, we provided two anecdotal stories about the *stringent clearance* during January 2020.

It is noteworthy that the 14-day observation window has already been set as epidemiological criteria to define a suspected case since January 18, 2020 (92) and was publicly mentioned in a National Health Commission guidance on January 22.[10] Following the central government's guidance, local governments soon adopoted this 14-day observation window as a decisive factor in their zero-Covid policies. This window will be a key instrument we use to construct our major treatment variable.

3.2.2 Second Phase: Stringent Clearance and Dynamic Clearance

Nearly one month after enforcing its stringent clearance policies,[11] the central government attempted to re-boost the economy and partially relax its public health interventions. On February 17, the State Council and National Health Commission of China issued *Prevention Guidance for Novel Coronavirus Pneumonia (version 5)* [12] which required local governments to classify different risk levels for different regions. Low risk areas, which are usually defined as prefectures with no COVID cases, should restrict travel from medium and high risk areas, while mobility within the prefecture and across other low risk areas were permitted. It is noteworthy that there still could be *dynamic*

[9]No additional prefectures adopted similar measures between February 20 to June 30, 2020 according to (115)

[10]NHC guidance: http://www.nhc.gov.cn/jkj/s3577/202001/c67cfe29ecf1470e8c7fc47d3b751e88.shtml.

[11]"In all Chinese cities, the Spring Festival holiday was extended, and people were advised to stay at home when possible, enforce social distancing and maintain good hygiene." (77)

[12]Prevention Guidance for Novel Coronavirus Pneumonia (version 5): http://www.nhc.gov.cn/jkj/s3577/202002/a5d6f7b8c48c451c87dba14889b30147.shtml.

clearance[13] policies implemented at low risk areas, such as school closings, cancellation of public events and restaurant closures. The medium risk areas were defined as prefectures without an *outbreak*.[14] On average, the high risk areas were defined as those with more than 10 cases reported within 14 days.[15] The medium and high risk regions were both subject to stringent clearance strategies, including traffic restriction, Fangcang hospital (mobile cabin hospital), community isolation and forced stay-at-home orders.[16] Although this state-issued *Guidance* left local governments with the freedom to manipulate the boundaries between high and medium risk levels, the high and medium risk areas could only become low risk after 14 consecutive days of no case increase. This is considered to be a clear distinction between low risk level and the other two levels [17].

Local governments immediately followed the central government's guidance. By the end of February, half of China's provinces were no longer at the First level reaction. It's possible that a Third level reaction province contains high or medium risk areas (prefectures), but the rest part of the province was more likely to adopt *dynamic clearance* policies or only keep travel restrictions between high risk areas. As of April 30, the national daily cases were already smaller than 50. Beijing and its neighboring provinces switched to the Second level. Three days later, Hubei switched to the Second emergency response level and no provinces remained at the First response level anymore.

[13] In 2020, "dynamic clearance" refers to implementing precise containment measures to control the spread of virus at small economic costs. However, this terminology was interpreted differently — to eliminate COVID at any cost — by Chinese propaganda in 2022, when Chinese government was dealing with Omicron variant. In this paper, we adhere to the definition of "dynamic clearance" from the year 2020.

[14] An *outbreak* is defined as 2 to 10 or more emerging confirmed COVID-19 cases within 14 days.

[15] The threshold between the medium risk and high risk were set quite differently across local governments

[16] Again, there is no general distinction between the clearance strategies for the medium and the high risk regions. In some cases, residents of high and medium risk regions were strictly required to stay at home, with security patrols checking on violators. Food and medicine could only be ordered through delivery

[17] Xinhua News Agency and People.cn (the two largest official state propaganda agencies of China) reported and confirmed that the 14 consecutive days of no case increase is used as the threshold for defining a low-risk level area.: http://www.gov.cn/xinwen/2020-03/23/content_5494361.htm and http://hn.people.com.cn/n2/2021/0815/c195194-34868198.html.

3.3 Data

3.3.1 CFPS Data

The individual data are from the China Family Panel Studies (CFPS), which is a nationally representative survey conducted by Peking University's Institute of Social Science Survey. This longitudinal survey covers 25 provincial-level regions in China (excluding Hong Kong, Macao, Taiwan, Xinjiang, Qinghai, Inner Mongolia, Ningxia, and Hainan), which accounted for 95% of China's total population.

We collect four waves of CFPS data, surveyed in 2014, 2016, 2018, and 2020, giving us a sample of 139,983 observations. To arrive at the sample used for analysis, we first exclude observations who (i) were surveyed by a proxy mode which lacks information on labor outcome (16,696 observations); or (ii) were full-time students (10,617 observations), resulting in a sample of 112,670 observations. We further restrict attention to individuals whose ages were between 16 and 64, and the sample size reduces to 93,357 observations. To keep consistency across main results and dynamic effect results, we drop respondents who were not interviewed in CFPS 2018, i.e., 17,141 observations. We drop 8,654 observations whose county is not included in the county list provided by Peking University's Institute of Social Science Survey in 2010. Finally, we drop 811 observations migrated to another county and 3,408 observations that appear only once in our sample. Finally, we end up with a sample of 63,343 observations (20,006 individuals). Among these 63,343 observations collected from 125 prefectures, 25.6 percent were surveyed in 2014, 27.6 percent were surveyed in 2016, 29.0 percent were surveyed in 2018 and the rest 17.8 percent were surveyed in 2020.[18] The main part of 2020 CFPS survey was conducted from July to August and a small share of interviewees were surveyed between September to December. There were mild COVID-19 spikes during this period and only 12 out of 126 prefectures in the CFPS sample were affected by the zero-COVID policy. This implies the zero-COVID policy is unlikely to cause systematic dropout of respondents and relieves the

[18]We also report the distribution of samples across four waves in Table A1.

concern about the selection bias problem of the survey. Furthermore, if there was any zero-COVID policy imposed on the interviewee's place of residence, the survey would be conducted via telephone.

Our main outcome variable concerns individual unemployment status. There are several questions related to employment status in the CFPS questionnaire. Specifically, interviewees (excluding full-time students) are asked for the following questions: (1) "Including agricultural work, waged job, self-employment and private business (housework and unpaid help do not count), have you worked for at least one hour last week?" (2) "Do you have a job, but you are currently on temporary vacation, sick leave or other vacation, or on-the-job training?" (3) "Will you return to the original job position in a certain period or within six months?" (4) "Are you running your own business which is currently in an off-season, but will resume after a while?" (5) "Is your agricultural work (including cropping, managing orchard, collecting agricultural and forestry products, fish farming, fishing, raising livestock, selling agricultural products in market, etc.) in an off-season?" If an interviewee answers "NO" to all questions above, the interviewee is unemployed; otherwise, the interviewee is on employment.[19] Specifically, the interviews in 2020 were conducted after the second half of 2020, ensuring that our sample can cover the zero-Covid policy shock.

Moreover, there is a question for employed interviewees rather than self-employed interviewees and business owners, "Including salary, bonus, cash benefit, material benefit, and excluding tax, insurances, and public housing, how much in total did you make from this job for the last 12 months?" We construct the outcome variable *Income* according to the answer to this question. Finally, the questionnaire has a question, "How many hours per week on average did you work for this job in the past 12 months?" We construct the outcome variable *Hours Worked* accordingly. To capture the responses of hours worked along the intensive and extensive margins, we also include unemployed workers and

[19] Again, the definition of unemployment we apply here is similar to U-4 unemployment, which includes unemployed and discouraged workers.

replace the missing values of hours worked with zero.

Panel A of Table 1 presents a statistic summary for labor outcomes in our sample. The average unemployment is 0.173.[20] Among employed workers, the average labor income is 20,992 RMB and the average hours worked per week is 46.3.

Furthermore, we calculate length of subsistence as the ratio between cash or deposit and family's yearly expenditure.[21] For families located at the bottom 20% income distribution who are extremely vulnerable to unemployment, their saving could only maintain their basic family expenditure for around 6 months.

To investigate the heterogeneous effects of COVID-19, we use a series of basic demographic information from CFPS 2018. Specifically, we report the heterogeneous effects for the following dimensions: gender, age, education and the age of the youngest child in the household. Panel B of Table 1 provides a statistic summary for these demographic characteristics.

3.3.2 zero-Covid Policy *Duration*

The *Duration* of zero-Covid policy implemented in each prefecture is our primary treatment variable. To document the days that a prefecture was labeled as a medium or high risk region, thus potentially the zero-Covid anti-contagious measures were implemented in the region, we manually collect the daily new COVID-19 cases from Jan 23 to June 30 by using the time-series data provided by *Dingxiangyuan* website, which collects the official daily release of COVID-19 cases at prefecture level.[22] [23] According to the national guidance for COVID-19 containment, a region will remain in medium or high risk level until a consecutive 14-day without new confirmed COVID-19 case, then the risk

[20] In Table A2, we report the different measures of unemployment in China and the United States. One could observe the U-4 unemployment rate collected from CFPS is higher than the official U-3 unemployment rate published by the Chinese government. The discrepancy between these two measures is also larger in China than in US.

[21] Figure A2 displays how many years interviewees' cash or deposit could afford their expenditure if they become unemployed and have no other income.

[22] Jan 23 was the time point when Wuhan lockdown and provinces enacted First Level emergency response.

[23] CFPS 2020 survey was collected during the second half of 2020. We would like to ensure the surveyed individual was exposed to the influence of the zero-Covid policy and the pandemic before taking the survey.

level will degrade to low. We measure each prefecture's medium or high risk period by excluding the low-risk period, i.e., the dates that have no COVID-19 cases and are not within a 14-day window of new COVID-19 case. Essentially, *Duration* measures how many days that a prefecture was exposed to mid- or high- risk under the national 14-day observation rule, accompanied by a wide spectrum of anti-contagious measures under the zero-Covid policy. Panel C of Table 1 summarizes the statistics for zero-Covid policy duration and COVID-19 cases at prefecture level. Average zero-Covid policy duration is 37.128 days. Average number of confirmed cases and death is 451.697 and 31.432, respectively. 34.9 percent of the prefectures once implemented a (city level) lockdown policy. Finally, for regression estimation, we use *lnDuration*, the log of zero-COVID policy duration plus 1, as the major treatment variable.

One possible concern is that our measure of zero-Covid policy duration is constructed from the COVID-19 case data following the guidance rule enforced by the central government, instead of documenting the real duration of mid- or high- risk level in each prefecture. To test the validity of the treatment, the ideal way is to compare the documented prefecture-level zero-Covid policy duration from January to June, 2020 with our constructed treatment. However, there is no accurate measure for the timing and duration of the zero-Covid policy corresponding to this period.[24] (77) and (136) collect information on starting dates for lockdowns without ending dates, and thereby cannot provide accurate duration of lockdowns. (74) generate stringent indices for China's COVID-19 responses, however, the policy stringency is measured at the province level. To our best knowledge, (33) and (63) are the only two studies that provide the timing and duration of lockdown policies and risk level indicators in China. Nevertheless, (33)'s research collects data between April 2020 and January 2022, only 2 months overlapped with the period considered in this paper and (63)'s research collects risk level indicators since April 2021, which is not overlapped with this study.

[24] The uniform national rule has not been launched by April 2020. Pre-April, it is hard to classify different local risk level rules into a general framework. For example, Zhejiang province use a five color system to classify risk level before April 2020.

We, therefore, choose an alternative outcome, the traffic mobility index from Baidu (China's Google mobility index) (83), to validate our treatment. To be specific, we calculate the difference between the daily traffic mobility indices (including in-town traffic, out-town traffic, and intra-town traffic) in 2020 and those of a comparable lunar date a year ago in each prefecture. [25] [26] We define *Exposed to Risk* as a dummy that is 1 if prefecture p is considered as mid- or high- risk on date t, under the national 14-day observation rule. We expect to observe a significant negative correlation between traffic mobility and the zero-Covid policy indicator. The period of traffic index we use is from January 1st to May 7th, which has 4 months overlapped with the period considered in our sample. [27]

As shown in Equation (1), we further apply a DiD setting to examine the validity of our treatment — *LnDuration*, the log form *Duration* of the instrumented zero-Covid policy. ΔY_{pt} is the measure of difference in mobility for prefecture p on date t. Besides, *Jan23* is a dummy that is 1 for the dates post January 23, 2020, and θ_p and δ_t capture the prefecture and date fixed effects.

$$\Delta Y_{pt} = \beta(lnDuration_p \times Jan23_t) + \theta_p + \delta_t + \epsilon_{pt} \tag{3.1}$$

We report the correlation between *Exposed to Risk* and change in traffic mobility relative to 2019 in the first three columns of Table A3. In Panel A, among all the prefectures in China, columns (1) to (3) show negative and significant correlations between whether exposed to zero-Covid policy and the decline of in-town, out-town, and intra-town traffic mobility. Columns (4) to (6) present the DiD estimation results which suggest a non-trivial decline in traffic mobility relative to the counterfactual change in mobility based on 2019 corresponding to the duration of zero-Covid exposure. In Panel B, we only

[25] Since the research period is overlapped with the Chinese Spring Festival, involving with high volatility in traffic mobility, we use the Chinese lunar calendar as a comparable date

[26] A similar calculation was used in (118)

[27] Baidu stopped publishing the mobility index after May 8th, 2020.

keep the CFPS surveyed prefectures in the sample, and the results stay robust. These stable significant negative correlations provide evidence for the validity of our treatment variable — as a prefecture receives a larger treatment, the more it is possibly exposed to the zero-Covid policy.

Another concern is that local officials would not implement corresponding zero Covid policies even if the prefecture is listed as medium or high risk level. After China's government admitted the outbreak of COVID, "Zero Covid" was put at the top of the list and is considered as an important index for local officials' performance evaluation. Officials who failed to contain COVID virus, responded slowly to new cases, or implemented inadequate prevention policies would be punished or dismissed. [28] Furthermore, the State Council encouraged people to directly report to the central government if local officials were considered as passive with zero Covid policy. [29] As a consequence, the "zero Covid" has become a Mao-style political campaign that local officials have strong incentives to immediately react with risk level indicator and therefore implement related policies, sometimes even over-reacted. [30]

It is important to point out that our measure also captures less stringent zero-Covid interventions other than lockdowns. [31] As we argued in Section 2, for prefectures with mild increases in COVID-19 cases, less stringent policies are more likely to be implemented as they are enough to mitigate the spread of the virus. In Figure A3, we plot the number of confirmed COVID-19 cases versus zero-Covid policy duration for each prefecture, while categorized by whether prefectures experienced lockdown or not.[32] We could observe that

[28] China dismisses 5 officials in Tibet for "inadequate" COVID prevention work: https://savetibet.org/china-dismisses-5-officials-in-tibet-for-inadequate-covid-prevention-work/; Officials from Yunnan and Hubei provinces were punished: http://www.xinhuanet.com/politics/2020-04/27/c_1125914123.htm

[29] Online complaint platform: http://www.gov.cn/xinwen/2020-01/24/content_5472009.htm

[30] See the essay from New York Times: https://www.nytimes.com/2022/04/13/business/china-covid-zero-shanghai.html?_ga=2.228845978.782308438.1670148502-1980932841.1670148502

[31] Lockdown is classified as the most stringent policy by (7).

[32] The lockdown information is adopted from (77). They defined a city (prefecture) implementing lockdown "when the following three measures were all enforced: (1) prohibition of unnecessary commercial activities for people's daily lives, (2) prohibition of any type of gathering by residents, (3)restrictions on private (vehicles) and public transportation."

prefectures with similar situations in COVID-19 cases and zero-Covid duration could vary in their lockdown decisions, which implies that a dummy variable for lockdown could not fully capture the spectrum of zero-COVID policies that a prefecture implemented.

The last potential concern about the construction of our treatment variable is that the intensity and the coverage of the anti-contagious measures during the early stage of the virus outbreak were more stringent compared to the later period when dynamic clearance was recommended by the central government. To cope with this issue, we further construct two separate duration variables corresponding to different time periods (using the same method described above): one for the period between Jan 23 and Feb 17; another for the period after Feb 17 till the starting date of survey collection (June 30).[33] In this way, we are able to capture the effects of zero-Covid policies on labor market outcomes in different phases of the pandemic.

3.3.3 Prefecture-Level Data

In addition to COVID-19 case data, our empirical analysis relies on other prefecture-level data that come from the 2018 China City Statistical Yearbook. These variables include (1)Population; (2) Gross Domestic Product (GDP); (3) Share of Service Sector in GDP. Panel D of Table 1 summarizes statistics for prefecture characteristics in 2018. The average population is 5.586 million and the average GDP is 396.489 billion RMB[34]. In addition, we provide prefecture government evaluation scores that have been aggregated from individual levels. The original score range is from 1 to 5, with 1 representing the belief that the local government achieved great success and 5 representing the belief that the local government achieved nothing.

[33] On Feb 17, State Council issued an official document that regions should be classified into three different risk levels, as a plan to boost the economy

[34] The minimum GDP is 15937.7 thousand RMB and round to 0 in billion.

3.4 Identification

3.4.1 Baseline Model

We begin by examining whether the zero-Covid policy in China induces an increase in individual-level unemployment probability by estimating a generalized Difference in Differences model:

$$Y_{ipt} = \beta(lnDuration_p \times Post_t) + \sum_{t \in \{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt} \tag{3.2}$$

where Y_{ipt} represents the outcomes of interests (e.g., unemployment and working hours) of individual i, in prefecture p surveyed in year t. $lnDuration_p$ is constructed by the method mentioned in Section 3.2, which measures the duration of the zero-Covid policy at prefecture p in 2020 in log form. $Post_t$ is an indicator function that assigns one if the observation is from the treated year 2020 and zero otherwise.

The parameter of interest β captures the marginal effects of exposure to zero-Covid policies on labor outcomes. In contrast with binary treatment DiD, the continuous treatment captures more variation in the data, the marginal effect provides more policy implication in the real world and allows for comparative discussion with evidence from other countries (73; 12; 13). For robustness purposes, we also generate a binary treatment variable that assigns one if the $lnDuration_p$ is above the median.[35] We will discuss more details about the continuous treatment setting and potential challenges in Section 5.4.4.

To allow time-invariant individual characteristics to influence unemployment or hours worked, we include individual fixed effects, θ_i. To absorb trends differing across provinces, we include province-year fixed effects, $\delta_{r,t}$. $Year_t$ is a series of binary indicators for year 2014, 2016, and 2020, and the dummy for year 2018 (t=3) is omitted in the equation. X_p is a set of proxies for prefecture economic status, including population, GDP, and share of

[35] We split the sample at the median prefecture so that the number of treated and controlled prefectures is approximately balanced. This method is used by (86)

service industry in 2018. We include $X_p \times Year_t$ to let their effects differ across years, and thereby to address the concern that prefectures with different economic characteristics may respond differently to the pandemic through other channels.[36] In addition, we further control time-varying government evaluation scores in order to isolate the treatment effect from the effect of government ability. We cluster standard errors at prefecture level. In addition to the baseline setting, we use alternative clustering choices (province level, prefecture-year two-way clustering, province-year two-way clustering) as robustness checks.

3.4.2 Dynamic Model

Similar to (73), our generalized DiD design relies on the assumption that after adjusting for controls and fixed effects, the patterns in outcome variables would follow a common path in the absence of zero-Covid policy. We employ a dynamic model to examine this assumption.

$$Y_{ipt} = \sum_{t\in\{1,2,4\}} \beta_t(lnDuration_p \times Year_t) + \sum_{t\in\{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt} \quad (3.3)$$

In this model, the parameter of interest β_4 represents the relative effect of the duration of the zero-Covid policy. β_1 and β_2 provide the estimates of the relative impact on labor market outcomes up to six years prior to actual treatment. If the common path assumption holds, we should not observe a significant relative impact from the "placebo" treatments on the pre-treated outcomes. Same with the previous section, we also use a binary treatment DiD setting as a robustness check. In this case, the underlying assumption is the common trends in pre-intervention outcomes between treated and control groups.

[36] A detailed discussion of this method is made by (86). One possible alternative is to control time-varying characteristics, which means that post-treatment variables will be included in the regression. However, It will result in a "Bad Control" problem (8). Moreover, the data on prefecture level controls in 2020 is not available yet.

There are several threats to the identification assumptions underlying our generalized DiD design. First, the potential disproportionately distributed spillover effects from neighbor units would bias our estimation in either direction. Second, the anticipation of zero-Covid policy shock could impact post-treated outcomes through channels such as labor mobility or job opening. Third, the particular selection bias problem arises from continuous treatment DiD setting — heterogeneous gains across different treatment doses, given the same counterfactual treatment dose (27; 46). We discuss these threats to identification and present more evidence supporting our identification assumption in Section 5.4.

3.5 Results

3.5.1 Baseline Result

We first present our estimated zero-Covid policy effect on labor market outcomes using the baseline DiD specifications. In Table 2 Panel A, we provide estimates for the policy effct on unemployment. We denote our DiD estimator by the interaction of log form of the zero-Covid policy duration with an indicator for post-treatment. In column (1), we control individual fixed effects and year fixed effects, and the result suggests that a longer duration of the zero-Covid policy has a causal impact on the increased chance of unemployment. An 10% increase in the duration of the zero-Covid policy increases the individual unemployment probability by 0.08 on average, which is statistically significant at the 5% level. Since the local zero-Covid policies are largely determined by the provincial governments with unobserved factors that could be correlated to the spread of the pandemic, we include the province by year fixed effects in column (2). It is also possible for prefecture level factors to influence the local government's capacity in the anti-Covid campaign, so we control for the interaction terms of prefecture characteristics (log population, log GDP and share of service sector together) with year fixed effects in column (3) and (4), and we further control the government evaluation score in column (5). Our estimation of the average policy effects remains stable and statistically significant in all these specifications.

As our result estimates the impact of the policy duration based on a national rule, we are able to predict the counterfactual impact of anti-contagious policies with a shorter observation window. We construct the duration of zero-Covid policy in the counterfactual scenario where the required zero-Covid window reduces from 14 days to 5 days. Then we perform a back-of-the-envelope calculation and predict the policy effect on the labor market using the constructed data. We find that, compared to an increase of 0.0371 in the unemployment probability caused by the average policy duration of 37.128 days under current policy, the zero-Covid policy under a 5-day window would decrease the average policy duration to 30.096 days and only increase the unemployment probability by 0.0324, which is about a 12% decline in the marginal policy effect.

In Table 2 Panel B column (1) - (4), we present the estimated effects of the zero-Covid policy on log hours worked. It is noteworthy that our sample is limited to individuals who reported positive hours worked in 2020, so our estimates capture intensive margin responses. We find that the zero-Covid policy has a significant negative effect on the hours worked. A 10% increase in policy duration would reduce hours worked for the employed individuals by around 0.2% on average, equivalent to 0.1 hours per week, depending on the regression specification used and the results remain consistent with different controls. When compared to the average hours worked per week (46.54), the effect on hours worked appears to be small. However, there are several potential factors that could explain this result. First, the hours worked per week are calculated based on the past 12 months since the survey time (July to December, 2020), which includes both the pre- and post-COVID period. Therefore, the estimated policy effect on hours worked might be underestimated if the pandemic had a disproportionate impact on one part of the year. Additionally, our results may not capture the full policy impact due to spillover effects, which we will discuss in more detail in section 5.4.1 where we provide estimation of the policy effect controlling for the spillover effect. Furthermore, the increase in unemployment might have led to an increase in workload for workers who remained employed, offsetting the negative effect on intensive margin responses. The small estimated effect could also suggest that the working schedule is relatively inflexible in response to the pandemic shock or that hours worked increased for workers who were quarantined at their workplaces (e.g. factories, hospitals, schools) but decreased for those who did not have access to their usual working places.

As a robustness check, we also estimate a binary treatment difference-in-differences (DiD) specification. To do this, we first divide the prefectures into high and low treatment groups, using the median value of the policy duration as the threshold. Then, we estimate the coefficient of the interaction term between the dummy variable for high treatment groups and the time indicator for post-treatment, using all the specifications considered in the baseline model.

The result of this binary DiD estimation are presented in Table 2 column (6). We find that on average, the probability of unemployment for individuals in the high treatment group is 0.028 higher compared to their peers in the low treatment group, at the significance level of 5%. The estimated effect on hours worked remains negative, but is more uncertain. As discussed above, there are several potential reasons for this. Overall, the binary DiD estimation provides additional evidence of the policy effect on unemployment status and suggests a more complex patterns for hours worked.

We also report the estimated impact of the policy on the number of hours worked for the entire population, including those who did not work at all. Naturally, in the continuous DiD settings, we see that the policy's effect on hours worked is larger than in the intensive margin responses. Furthermore, in the binary DiD setting, the policy is associated with a significant decrease in hours worked at the 5% level, indicating that it may have reduced the number of hours worked for unemployed individuals .

We report the estimated impact of the zero-COVID policy on the log labor income of individuals who reported positive earnings. Our results indicate that a 10% longer policy duration could result in a 2% decrease in income, after controlling for various factors. When only controlling for individual fixed effects, the magnitude of the negative policy effect decreases to around 1% and becomes statistically insignificant, which implies that the policy effect on labor income is correlated with the regional factors. In column (5), the coefficient for the binary treatment is also statistically insignificant, potentially due to the same reasoning as the hours worked result.

3.5.2 Dynamic Effects

The DiD estimator is based on the assumption that prefectures with different policy duration would have parallel trends in employment before the implementation of the policy. The observed increase in unemployment probability is assumed to be due solely to the pandemic containment measures, not any unobserved prefecture characteristics associated with the pandemic outbreak. To test for potential violations of the parallel trend assumptions of the Did framework, we estimate the effect of the interaction terms between policy duration and the dummy variable for all survey years on unemployment. Figure 1 reports the estimated dynamic effect result. We observe that before the pandemic shock in 2020, prefectures with longer policy duration had no trend in unemployment. The estimated coefficients for 2014 and 2016 were not statistically different from zero, with 2018 as the base year. Only the coefficient for 2020 was positive and significant, indicating that the parallel trend assumption is likely to hold in our model. In Figure 2, we see similar results when examining the dynamic effect of the binary treatment variable. We report the dynamic effect estimation for the hours worked in Figure A4, which provides us a consistent pattern of parallel trends before the pandemic and a negative effect for 2020, although the significance disappears. This could be due to the rigidity in the working schedule for employed workers or other factors that we mentioned before.

3.5.3 Disentangled Effect

Disentangled from Health Effect

China provides a suitable empirical setting to investigate the sole impact of the anti-contagious policies and our estimation presents the isolated effect of the zero-Covid policy on the labor market outcomes, without the influence of the public health shock. Our reasoning is that the pandemic was put under control very quickly after implementing the stringent disease preventive measures, thus there were few prefectures that experienced a considerable outbreak. By June 30, 2020, the total confirmed number in China was 83,534, around 50,000 cases were detected in Wuhan and another 18,000 cases were detected in Hubei province. As the number of confirmed cases is trivial compared to the prefecture population, the infection probability is close to zero and the workers should have no behavioral change during the period. Other evidences ((97; 131)) show that there is no significant change in excess mortality rate in 2020, which implies there is only limited indirect health shock caused by overwhelmed public health system or delayed treatment due to the mobility restriction policies during the pandemic. Given the large population base and the relatively small number of confirmed cases, the health effect of the pandemic was arguably negligible in most parts of China.

However, the outcomes of interests could still be affected through psychological channel — at the beginning of the pandemic, people had limited knowledge to the virus and might choose to stay at home voluntarily for safety concerns. The first few confirmed, or death cases emerged in the region could still generate a psychological shock to the people and disturb the local market. To distinguish between the policy effect and the health effect, we apply the approach used by (106). We have already shown that the disruption on the labor market comes from the zero-Covid policy. This relationship could be due in part to the adverse effects of the health shock on the labor market outcomes. We account for this by controlling directly for confirmed cases and death cases. In equation (4), $lnCases_p$ is the prefecture level total confirmed cases in log form. $lnDeaths_p$ is the prefecture level confirmed death cases in log form. Both variables are counted between Jan 23 to June 30, 2020.

$$
\begin{aligned}
Y_{ipt} = &\beta(lnDuration_p \times Post_t) + \omega_1(lnCases_p \times Post_t) + \omega_2(lnDeaths_p \times Post_t) \\
&+ \sum_{t\in\{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt}
\end{aligned}
\tag{3.4}
$$

We estimate the DiD treatment effect of the number of confirmed cases and dead cases and report the results in Table 3. In column (1) (2) and (3), besides the DiD treatment for the policy duration and other standard fixed effects, we further include the DiD treatments for the public health shock, which are the interaction terms between the dummy variable for 2020 and number of confirmed cases, number of death cases and both in the regression, respectively. The results show that none of the public health shock estimators are positive or statistically significant, while the coefficient for policy duration does not change much. This implies that the potential public health shock did not measurably influence the local employment status as the strict containment policy eliminated public health concerns efficiently. In other words, the results suggest that our estimated policy effect is not driven by the public health shock, but rather majorly reflects the impact of the zero-Covid policy on the labor market.

Disentangled from Lockdown Effect

As previously mentioned in Section 2, while economic activities were permitted in low risk areas, the policies implemented in mid and high risk areas were not explicitly outlined by the central government. This lack of clarity has allowed for different approaches by local governments, with some implementing flexible measures to support economic recovery in the mid risk area, while others implemented extremely strict measures to prevent the spread of the virus.

To confirm that the policy effect is not majorly driven by these stringent measures, e.g. prefecture-level lockdowns, implemented by local governments during the early stage of the pandemic, and disentangle the effect of policy intensity and policy duration, we include indicator variables that track whether the prefectures have ever implemented lockdowns during our sample period. These lockdowns are categorized as prefecture-level and community-level by (77). The prefecture-level lockdowns are defined as inter-city travel restriction, and the community-level lockdowns are defined as intra-city mobility restriction. It is noteworthy that our treatments of policy duration additionally capture low-intensity containment measures neglected by the lockdown variable. For example, for a prefecture that never issued within or between cities mobility restriction, it may still have issued a stay-at-home order fora specific district or area that is potentially exposed with COVID-19 cases. In the following model, parameter π_1 and π_2 absorb the lockdown effect and isolate β as the effect generated from the duration of the general disease preventive policy.

$$\begin{aligned} Y_{ipt} &= \beta(Duration_p \times Post_t) + \pi_1(Lockdown\_city_p \times Post_t) \\ &\quad + \pi_2(Lockdown\_comm_p \times Post_t) + \sum_{t \in \{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt} \end{aligned} \tag{3.5}$$

In Table 4 columns (1)(2) and (5)(6), we estimate the DiD treatment effects of the zero-Covid policy on individual unemployment status and hours worked controlling for the lockdown variables. We include the interaction term of the dummy variables for lockdown and the dummy variable for 2020 in the baseline regression to test whether lockdown is the major driven factor of the labor market disruption. In these regression specifications, the estimators of the policy duration remain statistically significant and the magnitude of the coefficients are similar to the baseline results. In columns (3)(4) and (7)(8), we only estimate the effects of the DiD treatment for lockdown variables solely on the labor outcomes and the coefficients are all statistically insignificant. These results imply that whether a city implemented lockdown could not fully explain the negative pattern observed in the labor market. We provide further evidences that the zero-Covid policy, disentangled from the city lockdown, made a causal impact on the labor outcomes.

3.5.4 Threats to Baseline Findings

Spillover Effects

Our baseline estimation relies on the assumption that the prefectures in our sample were not affected by the anti-contagious policies of neighboring prefectures. Potentially, the labor market is not only affected by local anti-contagious policy, but also be influenced by spillover effects from nearby regions. The inter-region traffic and human mobility could be strictly controlled, and therefore decreases local working opportunities. If the zero-COVID spillover effect disproportionately drove up the unemployment probability between sample prefectures, our estimation could be biased.

For example, if there exist stronger spillover effects (impacted by neighbors) in prefectures with relatively longer zero-Covid policy duration, and weaker spillover effect (impacted by neighbors) in prefectures with relatively shorter duration, the coefficient of local policy effect will be overestimated. Alternatively, if prefectures with relatively shorter zero-Covid policy duration experienced severe spillover from neighbors and prefectures with relatively longer zero-Covid policy duration experienced negligible spillover effects, the policy impacts will be underestimated.

In this section, we empirically assess the *Stable Unit Treatment Values Assumption* (SUTVA) by controlling the duration of zero-Covid policy in nearby prefectures. If we observe a negative (positive) correlation between local labor outcomes and zero-COVID policy duration of nearby prefectures, it implies the estimates of local policy effect in the baseline model is overstated (understated) in magnitude.

To measure the duration of zero-Covid policy in nearby prefectures, we first collect the zero-Covid policy duration for all neighboring prefectures of the surveyed prefectures in our sample. Then, we define the $Duration\_Nearby_p$ as the average neighbors' policy duration for a given prefecture p.

$$Duration\_Nearby_p = \frac{\sum_q Duration_q * I(q,p)}{\sum_q I(q,p)}$$

where $I(q,p)$ is the indicator function for whether prefecture p and prefecture q are nearby.

Our estimation model for the policy effect controlling for spillover effects is following:

$$\begin{aligned} Y_{ipt} &= \beta(lnDuration_p \times Post_t) + \alpha(lnDuration\_Nearby_p \times Post_t) \\ &+ \sum_{t \in \{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt} \end{aligned} \tag{3.6}$$

In Table 5, we estimate the effects of both local policy duration and nearby policy duration on labor market outcomes. In column (1), we report the estimation of policy impact controlling for spillover effect on individual unemployment probability. The estimated local policy effect remains positive and statistically significant, while the spillover effect has a negative coefficient which is not significant. The coefficient for the local policy duration is also close to the estimates of policy effect in our baseline specification as shown in Table 2. These results imply that the spillovers are unlikely to be present as the nearby policy duration did not contribute to the increase, if not a decrease, in the individual unemployment probability.

In column (2), we report the estimation of policy impact controlling for spillover effect on log hours worked for employed workers. While the spillover effect is still not significant, the magnitude of the local policy effect on the decrease of log hours worked increases from 0.0239 to 0.0424, compared to our baseline estimation. The increase in the local policy impact and the positive coefficients for nearby policy duration suggest that our baseline model might underestimate the size of the negative impact of local policy on the hours worked. This result also helps understand that the trivial policy effect on hours worked in the baseline model could be due to the underestimation from spillover effect.

Anticipation

Another challenge to our identification strategy is that patterns of labor outcomes could change in anticipation of zero-Covid policy shock. Nevertheless, when the COVID-19 virus initially outbroke at China, the Chinese government did not admit that the coronavirus has human-to-human transmissibility until Jan 20, 2020. Three days later, Wuhan implemented the city lockdown as well as the whole nation started implementing stringent anti-contagious policies soon after. As the time interval between the outbreak and the roll out of unprecedented policies is so narrow for labor market to anticipate, it addresses the concern of pre-anticipation bias.

Placebo Test

We employ the method suggested by (85) as the placebo test. We use the truncated alternative version of the DiD model (drop the data in 2020 when the treatment actually happened) and choose 2016 or 2018 as fake treatment periods. We report our estimation results in Table A6. Since we cannot find significant policy effects at the fake treatment periods, it suggests that the common trends assumption is likely to hold and our baseline estimations are not contaminated by non-treatment influences.

In the literature, a falsification test is typically performed to examine the key identification assumption underlying DID, where treatment variation does not correlate with preexisting differences between treatment group and control group. However, the test can only address the concern that serial correlation can bias standard errors, leading to an over-rejection of the null hypothesis (41; 18). As we cluster standard errors at multiple levels, we don't need to implement the test.

Selection Bias

There are two sources of selection bias in the continuous DiD treatment setting — classic selection bias and differences in treatment effects across different treatment doses.[37] In Section 5.2 and 5.4.2, we have already resolved uncertainty on common trends, where is also referred as classic selection bias (46). In this section, we are going to discuss the later concern.

To identify causality with our continuous treatment DiD setting, we need a stronger parallel trends assumption that "for all doses, the average change in outcomes over time across all units if they had been assigned that amount of does is the same as the average change in outcomes over time for all units that experienced that dose" (27). If this assumption does not hold, the estimates will be biased. For example, among two prefectures with $lnDuration$ d_j and d_{j-1}, there might be heterogeneous policy effects at the same treatment level d_{j-1}, which will result in a selection bias. When we calculate the marginal policy effect, the selection bias is represented by the second term on the right

[37] According to recent discussion by (27),"Unlike classic selection bias which is the differences in Y(0) for two groups of people, the bias of a continuous treatment difference-in-differences comes from the heterogeneity in gains from the treatment. In other words, if groups of units have heterogeneous gains at some dosage, then the continuous treatment DiD is contaminated by differences in different dosage groups own expected returns."

hand side.

$$\frac{\partial \mathbb{E}[\Delta Y_t | D = d]}{\partial d} = \underbrace{ACRT(d|d)}_{\text{average causal responses}} + \underbrace{\left.\frac{\partial ATT(d|l)}{\partial l}\right|_{l=d'}}_{\text{selection bias}} \tag{3.7}$$

To be specific, our estimation might include Average Causal Responses (ACRT) and differences in Average Treatment Effect (ATT) across prefectures with differing $lnDuration$ at a given treatment level. Although there is no compelling method to assess the stronger parallel assumption mentioned above, we do not think the selection bias problem will seriously threat to our identification — the national level policy rule could alleviate the "select into different treatment dose" concern.

Given the number of days without 0 increase (absorbed by prefecture fixed effects), it is not easy for local governments to manipulate how many days with 0 new COVID-19 case in a 14-day window, as the time point of detecting a new case is quite random. We provide a hypothetical example in Figure A5: suppose there are two prefectures with identical characteristics and the period to calculate *Duration* is from Jan 23 to Feb 29, 2020. For prefecture A and B, the total confirmed cases are both 40 and the number of days with new cases are 13 and 14, respectively. The almost same pattern in COVID-19 cases should not be surprising because these two prefectures are comparable in all dimensions. In fact, the only difference is that for prefecture A, there are 3 cases confirmed on Feb 05, and for prefecture B, there are 2 cases confirmed on Feb 05 and 1 case confirmed on Feb 09. According to the national 14-day observation rule, the *Duration* (shaded area) for prefecture A is 28 (the start day of *Duration* is Jan 23, the end day is 14 days after Feb 05) and for prefecture B is 32 (the start day of *Duration* is Jan 23, the end day is 14 days after Feb 09). We believe that the last case in prefecture B detected on Feb 09 instead of Feb 05, is mainly driven by some random factors such as the COVID-19 testing turnaround time or the incubation period but not correlated with prefecture characteristics or manipulated by local government. The variation in the treatment is very likely orthogonal to “self selection”. Shown in Figure A6, conditional on the number of days with observed cases (X-axis),[38] we can observe large variation in our choice of treatment — *Duration* (Y-axis), which is driven by the random factors instead of prefecture characteristics.

The marginal treatment effect is less likely biased by the selection problem, given the fact that our treatment variable is exogenous conditional on prefecture fixed effects. However, there are still some chances that our baseline findings are influenced by outlier regions which are several largest prefectures that experienced severe lockdown or extremely long zero-Covid duration. Wuhan and Hubei province went through the initial COVID-19 outbreak and implemented stringent lockdown policies for the first two months of the pandemic. Big metropolises, including Beijing, Shanghai, Guangzhou, Chongqing and Tianjin, frequently detected new COVID-19 cases, resulting in very long zero-Covid policy duration. The treatment effect for individuals who live in these regions might be different from people living elsewhere, affecting the average treatment effect for the whole population. In Table A7, we report our estimation results excluding individuals who live in outlier regions. In columns (1) and (2), we drop individuals in Wuhan; in columns (3) and (4), we drop individuals in Hubei province, and in columns (5) and (6), we further drop individuals in big cities. The estimated policy effect on labor outcomes remain consistent and robust to the exclusions of these outlier regions.

3.5.5 Robustness Checks

Balanced Panel

As shown in Table A1, some of individuals did not answer the questionnaire for all four waves in our sample, and individuals' dropout condition might be influenced by the pandemic or some unobservable characteristics that are correlated to labor outcomes.

[38]which is correlated with prefecture factors and controlled by fixed effects in our econometric model

Based on this unbalanced panel, our estimation might be biased due to selection problems. As we argue in the section 3.1, the attrition is unlikely be correlated with the pandemic.

Furthermore, to ensure that our estimation is not dramatically influenced by the individuals' dropouts, we estimate the baseline regression specification for individuals. In Table A8, we only keep the individuals who stay in each wave from 2014 to 2020 and estimate the effect of zero-Covid policy on individual unemployment, hours worked and income with the remaining balanced panel data. Compared to the baseline result, the balanced panel estimations have a relatively larger magnitude in the coefficients with at least 10% significance level. This implies that the baseline estimation might underestimate the policy effect for labor outcomes, while the argument that there exists a causal impact of the zero-Covid policy on labor outcomes is not systematically challenged.

Cluster Robust

We want to confirm that our baseline statistical inference is not affected by alternative choices of clustering. In **??** columns (1) and (4), we re-estimate the baseline regression specification and implement the two-way clustering by prefecture and by year, allowing errors to be correlated across individuals within same prefecture and same year. In columns (2) and (5), we calculate the standard errors clustered at province level; in columns (3) and (6), we clustered the standard errors by province and by year. Although the standard errors become larger compared to our baseline specification, the statistical inferences on the policy effect are robust to different clustering methods.

Lag Effects

The 2020 CFPS survey took several months to collect the questionnaires across different regions in China. While the majority of the survey was collected during July and August 2020, a small share of the survey was collected later through the period from July to December 2020. The time variation in the data collection could potentially help us investigate whether the persistent policy impact on the local labor outcomes is varying in its lagging time.

We make the estimation for the subsample from each survey group whose questionnaires were collected in each month from July to December. In Figure A7, we report the coefficient and the standard error of the policy effect on unemployment estimated from the subsamples collected in each month from July to December. We could observe that the policy effect becomes insignificant as time goes, without clear trend of increasing or decreasing. Although this result is partially due to the sample size is smaller in the later month groups, it also implies that the impact of the zero-Covid policy on the unemployment does not have significant lag effect that is not captured by our major estimation. The survey data we use for our estimation result are still valid in analyzing the policy effect on labor outcomes in 2020.

Indirect Health Effect

When a region experiences a COVID-19 outbreak and its risk level increases to medium or high, the zero-COVID policy may restrict public access to hospitals to prevent the spread of infection among patients and medical staff. However, these measures can also limit the availability of public health resources for individuals with chronic disease who require ongoing medical treatment. This lack of access to treatment can lead to incapacitation and ultimately, unemployment for these individuals.

To investigate whether the impact of zero-COVID policies on labor markets was driven by the indirect health effect of a shortage of medical resources, we utilized the health information of respondents in the CFPS survey, specifically a dummy variable indicating whether the individual has a chronic disease. We included this variable in our DiD estimation and also estimated the main regression separately for groups of individuals who do or do not have chronic diseases. The results are reported in **??**. In column (2), we found that including the dummy variable for chronic disease slightly decreased the magnitude of the estimated policy effect. In columns (3) and (4), we found that the estimator for the policy effect remained about the same value for the group without chronic diseases, while the magnitude of the estimator more than doubled for the group with chronic diseases. These results indicate that within the population with a strong need for medical resources, there was a larger increase in the probability of unemployment due to the zero-COVID policy. However, since this group makes up a relatively small portion of the population (less than 10%), the overall impact of the zero-COVID policy on unemployment in the sample was not significantly driven by the indirect health effect of limited access to hospitals during the pandemic.

3.5.6 Heterogeneous Effects

Separate Phase: Stringent containment and Precise containment

As discussed in Section 2, the policy intensity during January and February is much stronger than the policy intensity after February. *Stringent clearance* measures, such as lockdown and stay at home order, are more likely to be rolled out between January and late February for pandemic containment purpose. After February 17, *dynamic clearance* measures, such as public place closing and travel restriction between risk areas, became prevalent. Although we cannot measure this granular intensity difference with available data, we use different phases, Jan to Feb 17 and Feb 18 to June, as proxies of *stringent clearance* and *dynamic clearance*.

In Equation (8), we use Feb 17 as cutoff for these two phases: Jan 23 to Feb 17, represented by $lnDuration\_JanFeb_p$ and Feb 17 to June 30, represented by $lnDuration\_FebJun_p$.[39] Between Jan 23 to Feb 17, each prefecture were under the province's First Level emergency response, with a smaller standard deviation in the treatment (shown in Figure A1). This estimation separates the policy effect for different phases of the pandemic, which policy implication will be discussed with more details later when we interpret the estimation results.

[39]Again, Feb 17 is the time point of central government guidance for *precise containment*

$$
\begin{aligned}
Y_{ipt} = {} & \beta(lnDuration\_JanFeb_p \times Post_t) + \eta(lnDuration\_FebJun_p \times Post_t) \\
& + \sum_{t \in \{1,2,4\}} (X_p \times Year_t)\lambda_t + \theta_i + \delta_{r,t} + \epsilon_{ipt}
\end{aligned}
\tag{3.8}
$$

As mentioned in Section 2 and Section 3.2, the zero-Covid policy in China experienced a shift around late Feb 2020. The central government issued a guidance to require the local governments to identify the areas exposed to the virus more precisely and limit the influence of the anti-contagious measures only in risky regions. While our estimation results indicate that the local policy duration cause a significant impact on labor outcomes, we are unsure that whether the intensity of the policy treatment is evenly distributed over the whole period from Jan 2020 to June 2020. Potentially, after the issue of the guidance in Feb, the intensity and extent of the zero-Covid policy is much restricted and the policy treatment effect is weakened compared to the early phase of the Covid pandemic.

To examine the policy effect during different time periods, we estimate the coefficients of the DiD treatment for policy duration before Feb 17, policy duration after Feb 17, and both of them, respectively. The results are reported in Table 6. Column (1) shows that the policy duration before Feb 17 is significant positively related to unemployment, while column (2) show that the policy duration after Feb 17 is not. The results keep consistent while we include both *stringent clearance* and *dynamic clearance* into regression, shown in column (3). Column (4), (5) and (6) display a similar pattern that there are only significant correlations between hours worked and *stringent clearance*, which implies that the magnitude of *dynamic clearance* after Feb 17 is less significant compared to the early phase.

Across-group

We estimate the heterogeneous impacts of policy duration on different sub-populations and the estimation results are shown in Table 7 and Table 8. We estimate the policy effect for different groups categorized by gender, age, education, income distribution rank and having a young child. The parameter of interest is the coefficient of the interaction term between the $lnDuration \times Post$ and sub-population indicators. We find that for groups such as female workers, workers above age 65, workers with education level less than middle school, the bottom 50 percent population in the income distribution, and parents whose children are younger than 6 years old, they are more vulnerable to the zero-Covid policy impact on the unemployment status, while whether they are employed by a private sector firm has no impact. Regarding the policy effect on employed workers' hours worked, none of these individual characteristics has an impact, potentially due to the fact that the rigidity in the working schedule limits the difference across different groups.

There could be also a potential labor outcome difference for workers in the agricultural sector versus non-agricultural workers. We re-estimate our baseline models for each group of works and report our results in **??**. We could find the non-agricultural workers experienced a stronger policy effect on their employment status than the agricultural workers, while the impacts on their hours worked are similar.

3.6 Conclusions

During the COVID-19 pandemic, countries across the globe adopted drastically different strategies to mitigate the unprecedented public health crisis. While China was the first to implement strict anti-contagious interventions nationally, the zero-Covid policy's impact on their economy remained obscure until recently. Based on a generalized DiD design, we find that when a prefecture's zero-Covid policy lasted for 10% (3.7 days) longer, on average, the individual-level unemployment probability increased by around 0.1, and employed workers lost 0.2% and 2% of their hours worked and income, respectively.

Our estimation disentangles the zero-Covid policy and the public health shock of COVID-19, with the latter having no significant effect on labor market outcomes. The impact of lockdown policies has been widely discussed in literature, while our paper focuses on the effect of China's zero-Covid policy, a full spectrum anti-contagious policy that includes not only lockdown but also less stringent anti-contagious measures that have been difficult to observe due to data limitations. We also considered spillover effects from nearby prefectures, which did not significantly contribute to the negative labor market impact. Additionally, our research suggests that only the stringent anti-contagious measures implemented during the early stage of the pandemic had a negative impact on labor outcomes.[40]

COVID-19 has caused millions of deaths and a global humanitarian crisis as many countries were unable to control the spread of the virus after the outbreak of the pandemic. Partially contributing to this catastrophic outcome is the fact that the potential economic and political outcomes of restricting human mobility deterred the policymakers from taking serious disease preventive measures immediately after the outbreak of the virus. We provide a systematic evaluation of the labor market disruption caused by the most stringent containment policy and estimate the economic cost of non-pharmacological interventions to stop the pandemic. It is noteworthy that the data used in this paper were collected during the period when the zero-Covid policy was very effective and the pandemic was controlled extremely well in China. It is reasonable to doubt that our estimation results are not valid under the circumstances where the spread of viruses is more difficult to put under control and the zero-Covid policy has to last longer.[41] The economic costs of the anti-contagious policy would not grow linearly as the

[40]Due to data limitation, we are not able to estimate the mid-term policy effect in 2021 — there was a strong rebound in the first half of 2021 and China's GDP growth rate reached 8.1 percent by the end of that year. It was partially credited to the zero-Covid policy in 2020. Also, we are not able to estimate the policy effect in 2022, when many regions in China experienced longer restrictions in mobility due to the outbreaks of the Omicron variant of COVID-19.

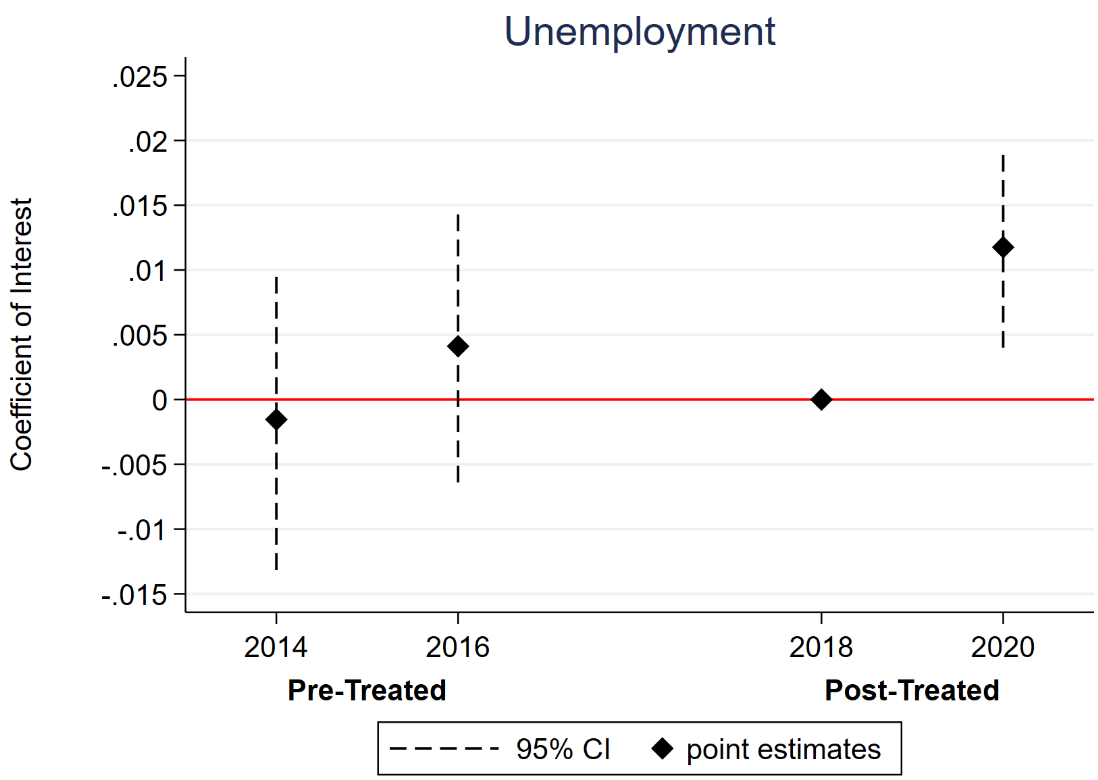
[41]As the mortality of COVID-19 has been reducing (129), the public health welfare brought by the zero-Covid policy has decreased dramatically. On Dec 7 2022, NHC announced a new guideline that loosened the zero-Covid restriction significantly.

duration of the policy increases, but exponentially.[42] However, our work can still serve as a benchmark under such a scenario: the pandemic's scope was constrained soon after its outbreak by fast and stringent containment measures, and millions of lives were saved. How much would it cost economically? After all, we hope our work will be a useful reference for future policymakers dealing with similar situations, where they will have to face the trade-off between health, freedom and economic well-being.

[42]Figure A3 displays how many years interviewees' cash or deposit could afford their expenditure if they become unemployed and have no other income. For families located at the bottom 20% income distribution who are extremely vulnerable to unemployment, their saving could only maintain their basic family expenditure for around 6 months.

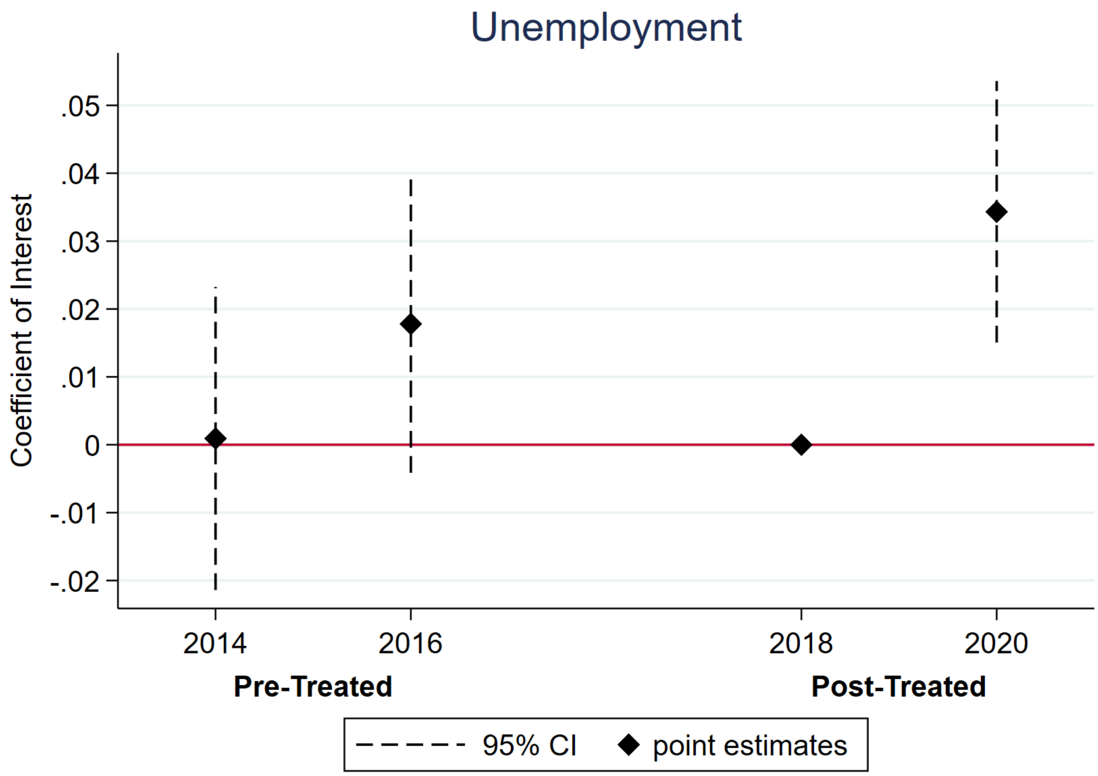
3.7 Figures and Tables

3.7.1 Figures



Notes: The figure shows coefficients and 95% confidence intervals from estimating the leads and lags regression in equation (3), where the dependent variable is unemployment. All effects are relative to 2018.

Figure 1: Dynamic Effects Unemployment (continuous treatment)



Notes: The figure shows coefficients and 95% confidence intervals from estimating the leads and lags regression in equation (3), where the dependent variable is unemployment. All effects are relative to 2018.

Figure 2: Dynamic Effects Unemployment (binary treatment)

3.7.2 Tables

	Obs	Mean	Std.Dev	Min	Max
Panel A: Individual D.V.					
Unemployed	63343	0.173	0.378	0.0	1
Hours Worked	37914	46.538	21.880	0.1	133
Hours Worked (Overall)	48601	36.003	27.247	0.0	133
Income	28445	20992.159	25262.527	0.0	100000
Panel B: Individual Characteristics					
Gender	63343	0.517	0.500	0.0	1
Age	63343	45.808	11.868	16.0	67
Education (middle school or below)	63343	0.732	0.443	0.0	1
Agricultural Worker	60215	0.432	0.495	0.0	1
Private Sector Worker	19669	0.841	0.366	0.0	1
Youngest Child Age	57690	19.120	11.394	0.0	47
Health Condition	63175	2.979	1.208	1.0	5
Chronic Disease	62771	0.152	0.359	0.0	1
Panel C: Prefecture Treatments					
Policy Duration	126	37.128	21.411	0.0	158
Policy Duration Feb Jun	126	18.349	19.158	0.0	135
Policy Duration Jan Feb	126	18.779	5.095	0.0	24
Covid Case Duration	126	13.921	13.054	0.0	102
Confirmed Cases	126	451.691	4481.225	0.0	50340
Confirmed Deaths	126	31.432	344.629	0.0	3869
Lockdown (City Level)	126	0.349	0.479	0.0	1
Lockdown (Community Level)	126	0.183	0.388	0.0	1
Panel D: Prefecture Controls					
Population 2018 (Thousand)	126	5586.448	4662.472	430.0	34040
GDP 2018 (Billion)	126	396.489	557.217	0.0	3268
Share of Service Sector in GDP	126	48.090	8.518	31.1	81
Government Evaluation	126	2.555	0.166	2.2	3

Notes: Panel A reports individual outcome variables of interest. Panel B reports descriptive individual characteristics. Panel C report prefecture-level treatment variables. Panel D reports prefecture-level characteristics in 2018.

Table 1: Statistic Summary

Table 2: Baseline Results: Unemployment and Hours Worked

	(1) Outcomes	(2) Outcomes	(3) Outcomes	(4) Outcomes	(5) Outcomes	(6) Outcomes
Panel A: Unemployment						
lnDuration × Post	0.00831**	0.0109**	0.0125***	0.0109**	0.0112**	
	(0.00385)	(0.00428)	(0.00443)	(0.00456)	(0.00463)	
Government Evaluation					-0.0227	
					(0.0154)	
lnDuration_Dummy × Post						0.0284***
						(0.00994)
R-squared	0.469	0.470	0.470	0.470	0.470	0.470
Observations	63343	63343	63343	63343	63343	63343
Mean of Unemployment	0.173	0.173	0.173	0.173	0.173	0.173
Panel B: log Hours Worked						
lnDuration × Post	-0.0165**	-0.0183***	-0.0215***	-0.0239***	-0.0245***	
	(0.00789)	(0.00642)	(0.00743)	(0.00764)	(0.00764)	
Government Evaluation					0.0587	
					(0.0733)	
lnDuration_Dummy × Post						-0.00793
						(0.0271)
R-squared	0.312	0.315	0.315	0.315	0.315	0.315
Observations	37914	37914	37914	37914	37914	37914
Mean of Hours Worked	46.54	46.54	46.54	46.54	46.54	46.54
Individual FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✗	✗	✗	✗	✗
Province-Year FE	✗	✓	✓	✓	✓	✓
lnPop × Year FE	✗	✗	✓	✓	✓	✓
lnGDP × Year FE	✗	✗	✗	✓	✓	✓
ServiceShare × Year FE	✗	✗	✗	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is an indicator for unemployment (Panel A) or the natural log of hours worked (Panel B). lnDuration is the natural log of the duration of the zero-Covid policy. lnDuration_Dummy is a binary variable that equals 1 if the duration of the zero-Covid policy is above median and 0 otherwise. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 3: Disentangled Effect: zero-Covid Policy and Public Health Shock

	(1)	(2)	(3)	(4)	(5)	(6)
	Unemployment	Unemployment	Unemployment	log Hours Worked	log Hours Worked	log Hours Worked
lnDuration × Post	0.0149**	0.0125**	0.0141*	-0.0357**	-0.0222**	-0.0399**
	(0.00726)	(0.00486)	(0.00754)	(0.0160)	(0.00892)	(0.0181)
lnCases × Post	-0.00363		-0.00182	0.0112		0.0204
	(0.00552)		(0.00736)	(0.0151)		(0.0230)
lnDeaths × Post		-0.00479	-0.00364		-0.00552	-0.0184
		(0.00614)	(0.00837)		(0.0156)	(0.0245)
R-squared	0.470	0.470	0.470	0.315	0.315	0.315
Observations	63343	63343	63343	37914	37914	37914
Mean of Outcome	0.173	0.173	0.173	46.54	46.54	46.54
Individual FE	✓	✓	✓	✓	✓	✓
Province-Year FE	✓	✓	✓	✓	✓	✓
lnPop × Year FE	✓	✓	✓	✓	✓	✓
lnGDP × Year FE	✓	✓	✓	✓	✓	✓
ServiceShare × Year FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

Unit of observation is an individual. The dependent variable is an indicator for unemployment (columns 1-3) or the natural log of hours worked (columns 4-6). lnDuration is the natural log of the duration of the zero-Covid policy. lnCases is the natural log of the number of confirmed cases. lnDeaths is the natural log of the number of death. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 4: Disentangled Effect: zero-Covid Policy and Lockdown

	(1) Unemployment	(2) Unemployment	(3) Unemployment	(4) Unemployment	(5) log Hours Worked	(6) log Hours Worked	(7) log Hours Worked	(8) log Hours Worked
Panel A: lnDuration								
lnDuration × Post	0.0110**	0.0105**			-0.0234***	-0.0256***		
	(0.00454)	(0.00475)			(0.00774)	(0.00833)		
Lockdown_city × Post	-0.00264		-0.00157		-0.0165		-0.0188	
	(0.0135)		(0.0135)		(0.0317)		(0.0316)	
Lockdown_comm × Post		0.00885		0.0117		0.0364		0.0296
		(0.0192)		(0.0193)		(0.0367)		(0.0378)
R-squared	0.470	0.470	0.470	0.470	0.315	0.315	0.315	0.315
Observations	63343	63343	63343	63343	37914	37914	37914	37914
Panel B: lnDuration_Dummy								
lnDuration_Dummy × Post	0.0285***	0.0279***			-0.00955	-0.00955		
	(0.00968)	(0.0103)			(0.0271)	(0.0263)		
Lockdown_city × Post	0.000548		-0.00157		-0.0197		-0.0188	
	(0.0130)		(0.0135)		(0.0320)		(0.0316)	
Lockdown_comm × Post		0.00884		0.0117		0.0305		0.0296
		(0.0178)		(0.0193)		(0.0375)		(0.0378)
R-squared	0.470	0.470	0.470	0.470	0.315	0.315	0.315	0.315
Observations	63343	63343	63343	63343	37914	37914	37914	37914
Mean of Outcome	0.173	0.173	0.173	0.173	46.54	46.54	46.54	46.54
Individual FE	✓	✓	✓	✓	✓	✓	✓	✓
Province-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
lnPop × Year FE	✓	✓	✓	✓	✓	✓	✓	✓
lnGDP × Year FE	✓	✓	✓	✓	✓	✓	✓	✓
ServiceShare × Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is an indicator for unemployment (columns 1-4) or the natural log of hours worked (columns 5-8). lnDuration is the natural log of the duration of the zero-Covid policy. lnDuration_Dummy is a binary variable that equals 1 if the duration of the zero-Covid policy is above median and 0 otherwise. Lockdown_city is an indicator that equals to 1 if the prefecture has ever experienced lockdown and 0 otherwise. Lockdown_comm is an indicator that equals to 1 if communities in the prefecture have ever experienced lockdown and 0 otherwise. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 5: Spillover Effect

	(1) Unemployment	(2) log Hours Worked
lnDuration × Post	0.0116***	-0.0253***
	(0.00437)	(0.00815)
lnDuration_nearby × Post	-0.0168	0.0327
	(0.0202)	(0.0332)
R-squared	0.470	0.315
Observations	63343	37914
Mean of Outcome	0.173	46.54
Individual FE	✓	✓
Province-Year FE	✓	✓
lnPop × Year FE	✓	✓
lnGDP × Year FE	✓	✓
ServiceShare × Year FE	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is an indicator for unemployment (columns 1) or the natural log of hours worked (columns 2). lnDuration is the natural log of the duration of the zero-Covid policy. lnDuration_nearby is the natural log of the average neighbors' policy duration for a given prefecture. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 6: Heterogeneity: Separate Phase

	(1) Unemployment	(2) Unemployment	(3) Unemployment	(4) log Hours Worked	(5) log Hours Worked	(6) log Hours Worked
lnDuration_JanFeb × Post	0.0146**		0.0130*	-0.0304***		-0.0267*
	(0.00579)		(0.00699)	(0.00823)		(0.0143)
lnDuration_FebJun × Post		0.00570	0.00203		-0.0125	-0.00473
		(0.00419)	(0.00471)		(0.0106)	(0.0131)
R-squared	0.470	0.470	0.470	0.315	0.315	0.315
Observations	63343	63343	63343	37914	37914	37914
Mean of Outcome	0.173	0.173	0.173	46.54	46.54	46.54
Individual FE				✓	✓	✓
Province-Year FE				✓	✓	✓
lnPop × Year FE				✓	✓	✓
lnGDP × Year FE				✓	✓	✓
ServiceShare × Year FE				✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is an indicator for unemployment (columns 1-3) or the natural log of hours worked (columns 4-6). lnDuration_JanFeb is the natural log of the duration of the zero-Covid policy between Jan 23 and Feb 17. lnDuration_FebJun is the natural log of the duration of the zero-Covid policy after Feb 17 until the starting date of survey collection (Jun 30). Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 7: Heterogeneity: Unemployment by individual characteristics

	(1) Gender	(2) Elder	(3) Middle School	(4) Private Sector	(5) Bottome Income Distribution	(6) Young Child
lnDuration × Post	0.00528	0.00489	0.000878	0.0147	-0.00428	0.0135**
	(0.00499)	(0.00497)	(0.00703)	(0.0180)	(0.00819)	(0.00614)
lnDuration × Post × Female	0.0106*					
	(0.00545)					
lnDuration × Post × Old		0.0122*				
		(0.00648)				
lnDuration × Post × edu_middle			0.0132*			
			(0.00789)			
lnDuration × Post × Private				0.00163		
				(0.0207)		
lnDuration × Post × income_bottom					0.0382***	
					(0.0114)	
lnDuration × Post × child_6						-0.0166*
						(0.00979)
lnDuration × Post × child_18						-0.0000539
						(0.00796)
R-squared	0.470	0.471	0.470	0.309	0.543	0.471
Observations	63343	63343	63343	19669	27651	63343
Mean of Unemployment	0.173	0.173	0.173	0.154	0.260	0.173
Individual FE	✓	✓	✓	✓	✓	✓
Province-Year FE	✓	✓	✓	✓	✓	✓
lnPop × Year FE	✓	✓	✓	✓	✓	✓
lnGDP × Year FE	✓	✓	✓	✓	✓	✓
ServiceShare × Year FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is an indicator for unemployment. lnDuration is the natural log of the duration of the zero-Covid policy. Female, Old, Edu_middle, Private, Income_bottom, Chind_6, and Child_18 are binary variables that equal to 1 if the interviewee is female, above age 65, with education level less than middle school, works in private firm, the bottom 50 percent population in the income distribution, has any child that is younger than 6 years old, has any child that is younger than 18 years old, and 0 otherwise. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 8: Heterogeneity: Hours Worked by individual characteristics

	(1) Gender	(2) Elder	(3) Middle School	(4) Private Sector	(5) Bottome Income Distribution	(6) Young Child
lnDuration × Post	-0.0219**	-0.0230**	-0.0385***	-0.0313	-0.0176	-0.0275**
	(0.00998)	(0.0107)	(0.0134)	(0.0294)	(0.0127)	(0.0115)
lnDuration × Post × Female	-0.00378					
	(0.0182)					
lnDuration × Post × Old		-0.00274				
		(0.0142)				
lnDuration × Post × edu_middle			0.0186			
			(0.0145)			
lnDuration × Post × Private				0.0164		
				(0.0282)		
lnDuration × Post × income_bottom					0.0231	
					(0.0288)	
lnDuration × Post × child_6						-0.000175
						(0.0208)
lnDuration × Post × child_18						0.0144
						(0.0242)
R-squared	0.315	0.316	0.315	0.213	0.231	0.316
Observations	37914	37914	37914	13536	16552	37914
Mean of Hours Worked	46.54	46.54	46.54	52.10	50.18	46.54
Individual FE	✓	✓	✓	✓	✓	✓
Province-Year FE	✓	✓	✓	✓	✓	✓
lnPop × Year FE	✓	✓	✓	✓	✓	✓
lnGDP × Year FE	✓	✓	✓	✓	✓	✓
ServiceShare × Year FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Unit of observation is an individual. The dependent variable is the natural log of hours worked. lnDuration is the natural log of the duration of the zero-Covid policy. Female, Old, Edu_middle, Private, Income_bottom, Chind_6, and Child_18 are binary variables that equal to 1 if the interviewee is female, above age 65, with education level less than middle school, works in private firm, the bottom 50 percent population in the income distribution, has any child that is younger than 6 years old, has any child that is younger than 18 years old, and 0 otherwise. Robust standard errors in parentheses, clustered at the prefecture level. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

3.8 Appendix

3.8.1 Anecdotal Evidence: Stringent Containment between Jan and Feb

Confronting a unprecedented public emergency case, Chinese local governments rolled out the most stringent containment policies during January to February, 2020. Although there is little detailed written instruction on how to conduct such containment policies, there were numerous news, coverage and videos on social media revealed local governments reaction by that time [43].

One suggestive example happened in Henan province. Although the daily increased cases is less than 50 and the rural regions were considered as the least affected areas, many villages blocked the entrance and do not allowed any form of visitors. In some cases, during Spring Festival, migrant workers who returned from work places were not allowed to enter the village. In one video on social media [44], village's Communist Party Secretary was using broadcast condemning a villager of hanging out, "are you even a human being? You are so fucked up", one of the public insults from the Secretary. Similar prefecture level or village level lockdown and traffic restrictions also launched in other parts of China (e.g. Heilongjiang, Zhejiang, Jiangxi, etc[45].), among the consequences, a truck driver's experience became a most ridiculous and black humorous story.

Mr.Xiao, a truck driver from Hubei, set off for Sichuan province since Januray 7. However, when he prepared the return trip on Januray 24, Hubei locked down. Mr Xiao had to drove away with no destination. The service areas refused him from stopping, the option of getting off the highway also became impossible, since all the cities rolled out travel restriction on people from Hubei. "People see my license plate, that I come from Hubei, and get scared". After seven days driving, he was found fall asleep in his truck on

[43] e.g. https://www.bilibili.com/video/BV1a7411k7NB?from=searchseid=5191564554814052769spm$_id_from =$ 333.337.0.0; https://www.bilibili.com/video/BV1H7411g75d?from=searchseid=5191564554814052769spm$_id_from =$ 333.337.0.0; https://www.bilibili.com/video/BV1n7411W7uH?from=searchseid=5191564554814052769spm$_id_from =$ 333.337.0.0; https://www.tuliu.com/read-121860.html;

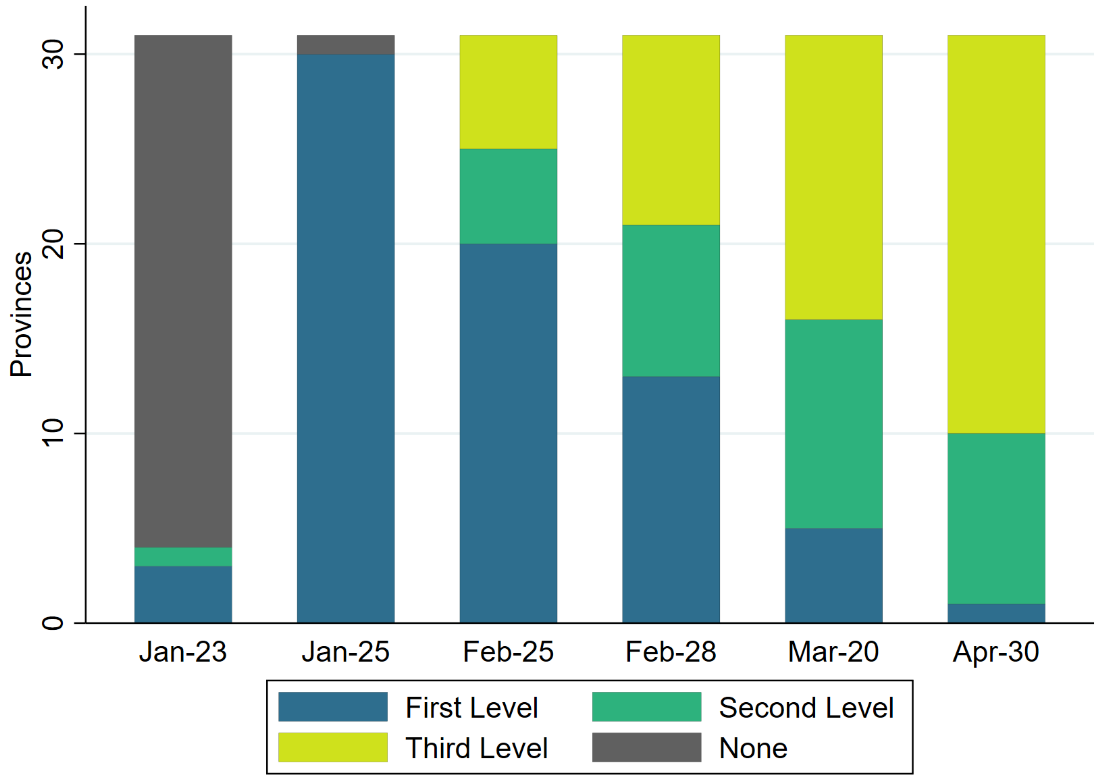
[44] Source: https://www.bilibili.com/video/BV1Y741167Yp/?spm$_id_from = autoNext$.

[45] Source: http://www.moa.gov.cn/xw/qg/202002/t20200224$_6$337603.htm.

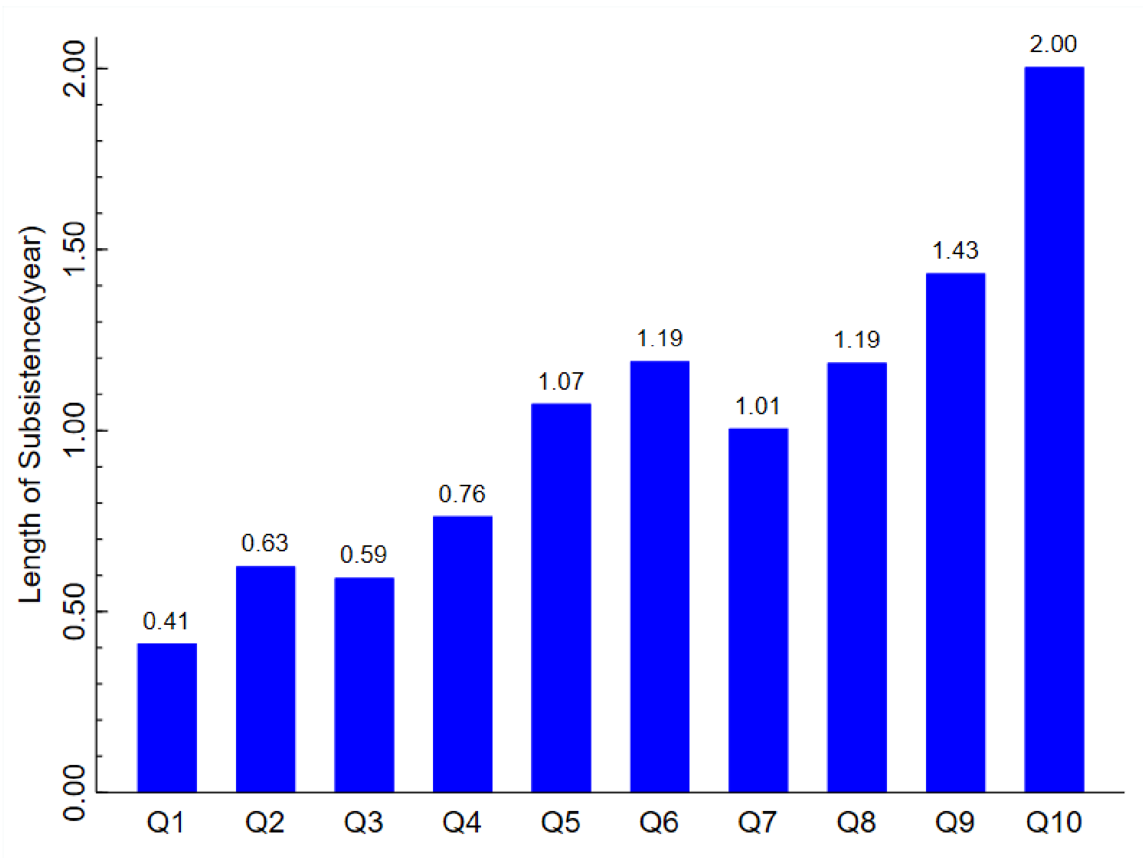
the emergency lane in Shaanxi province, thousands miles away from his home, “my greatest hope is that I can find a place to stop, get some good sleep and eat something.”. Fortunately, police officers got him a hotel room in a service area, Mr Xiao returned back home on Mar 16, 68 days after his adventure [46].

[46] Source: https://news.cgtn.com/news/2020-02-10/The-road-back-to-Hubei-Truck-driver-says-long-journey-still-not-over-NY4ba2qOaY/index.html; https://www.wsj.com/video/truck-driver-stuck-on-highway-since-chinas-coronavirus-lockdown/F7097DE9-FCEB-4E13-BCAB-9D715AF84D0B.html; https://www.sohu.com/a/371766675_617717.

3.8.2 Appendix Figures

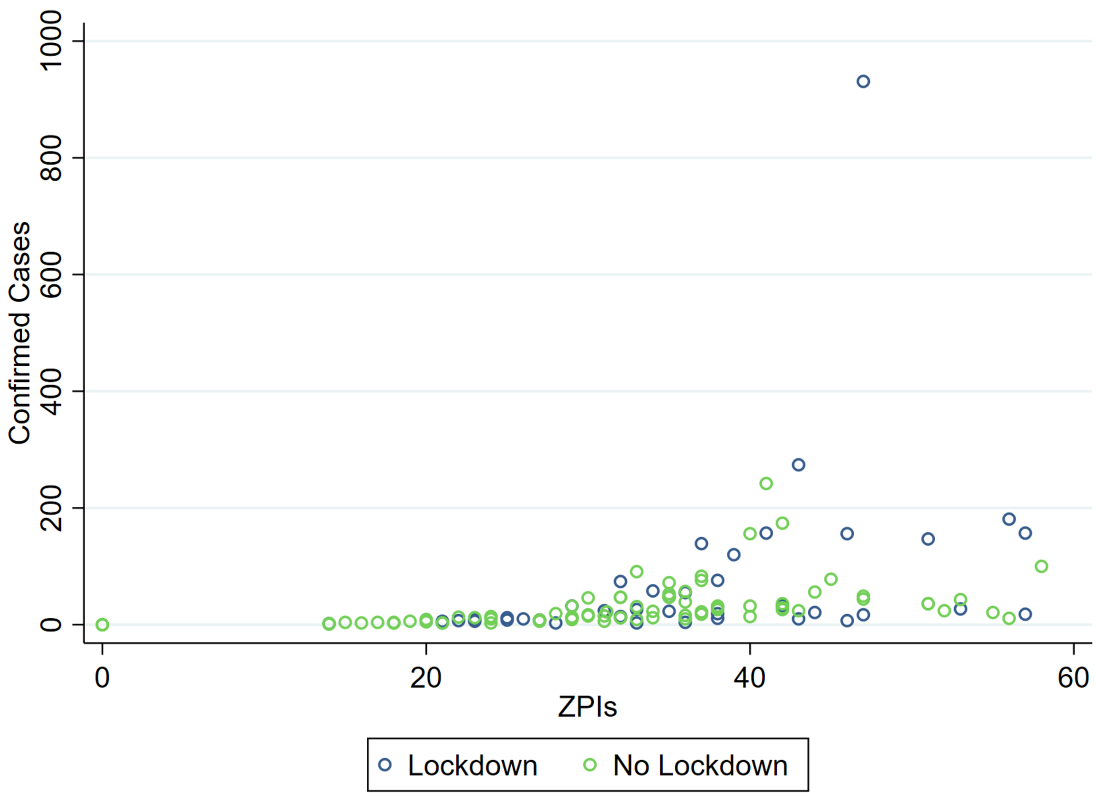


Notes: The figure shows the timeline of the provincial emergency reaction level.
Figure A1: Province Emergency Reaction Level Time Line



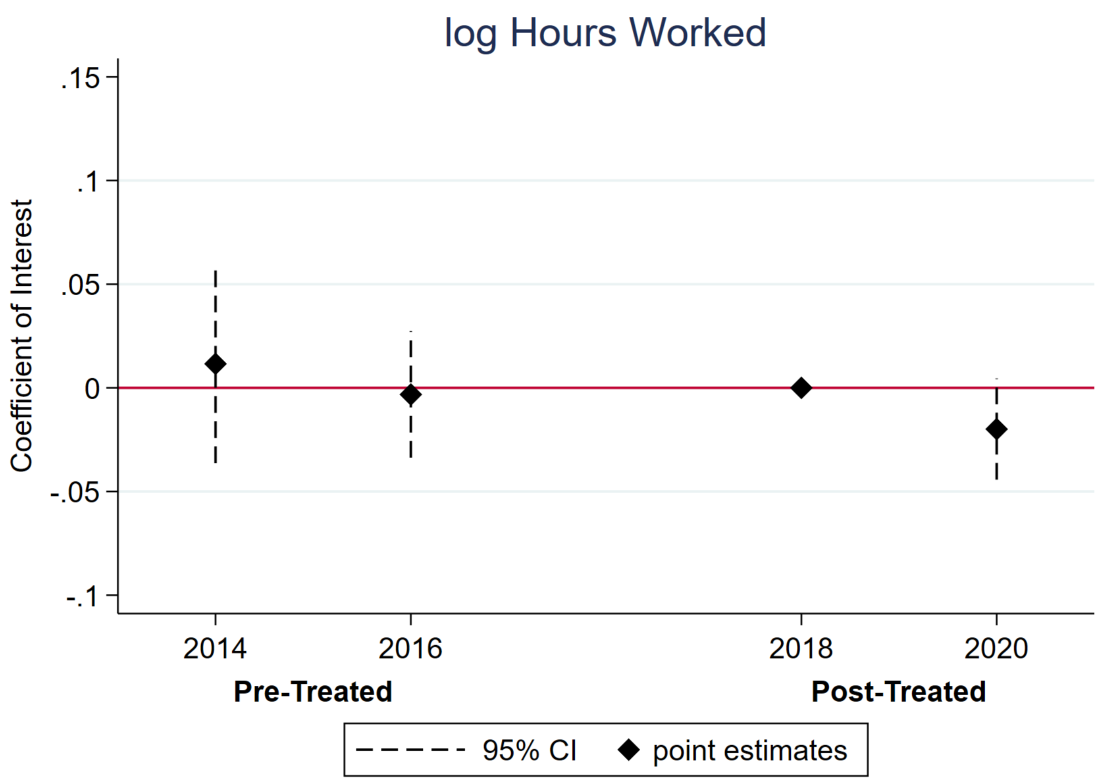
Notes: The figure shows how many years interviewees' cash or deposit could subsist their expenditure if they become unemployed. Y-axis represents length of subsistence = cash or deposit/ family's expenditure. X-axis represents deciles at income distribution.

Figure A2: Subsistence Years After Unemployed



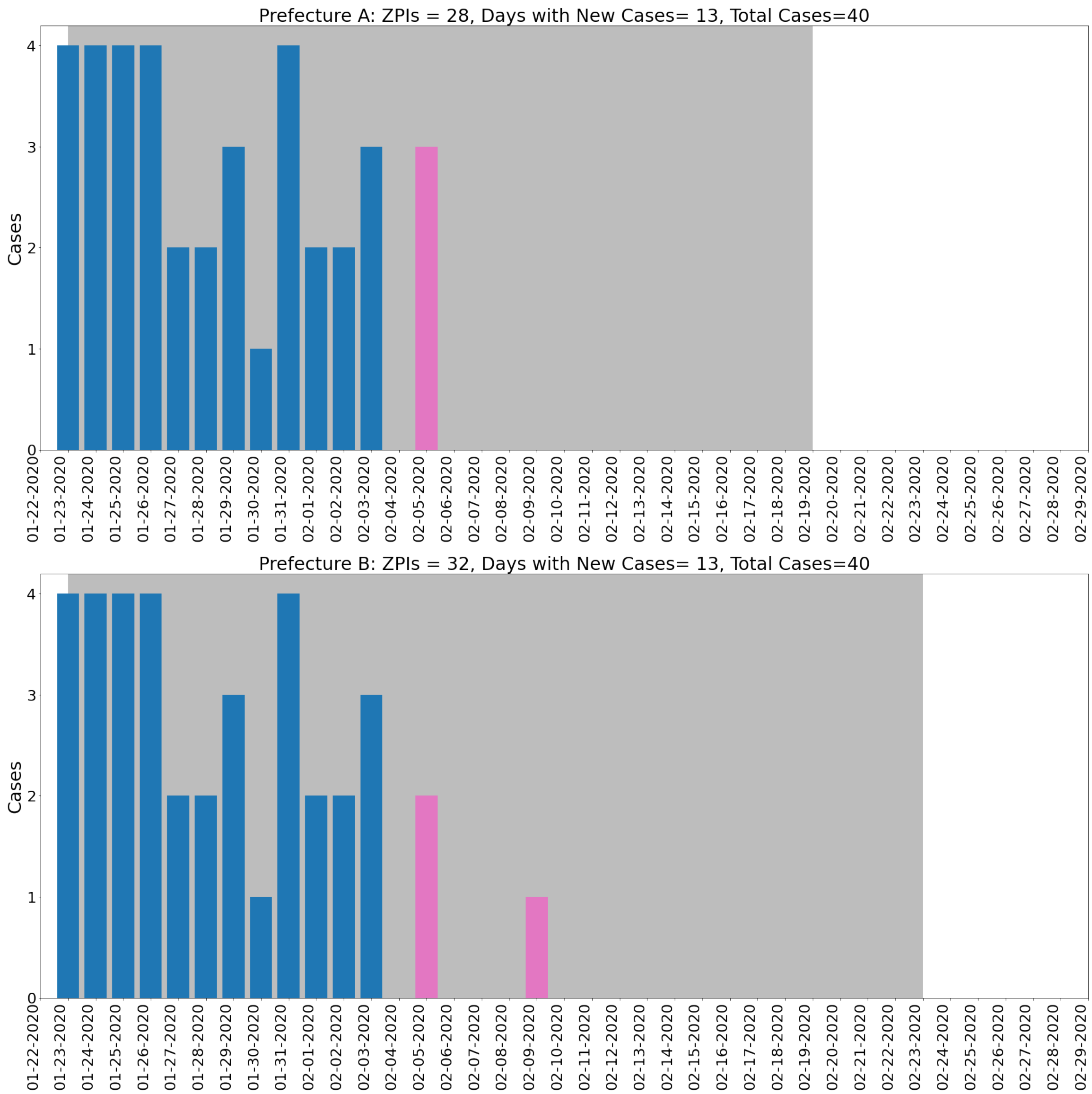
Notes: The figure shows the number of confirmed cases and the zero-Covid duration for prefectures by lockdown status. Duration outliers (95 percentile) are dropped from this graph.

Figure A3: Duration and Confirmed Cases by Lockdown



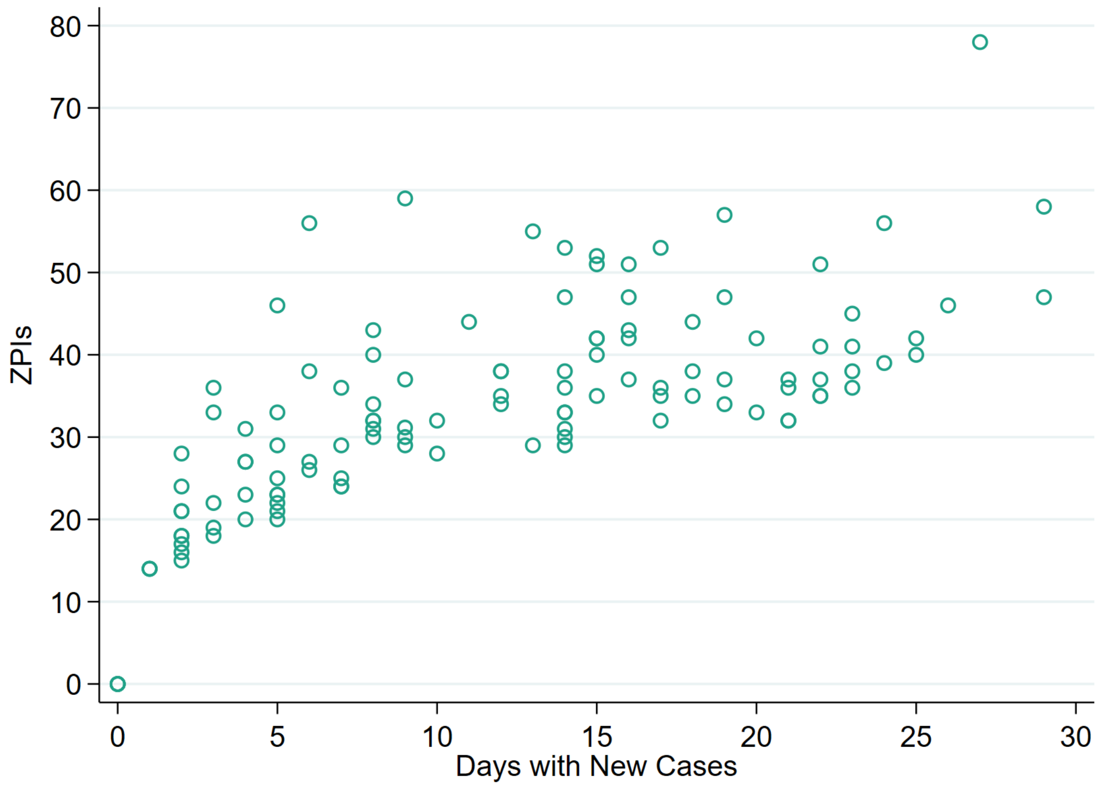
Notes: The figure shows coefficients and 95% confidence intervals from estimating the leads and lags regression in equation (2), where the dependent variable is the natural log of hours worked. All effects are relative to 2018.

Figure A4: Dynamic Effects: Hours Worked



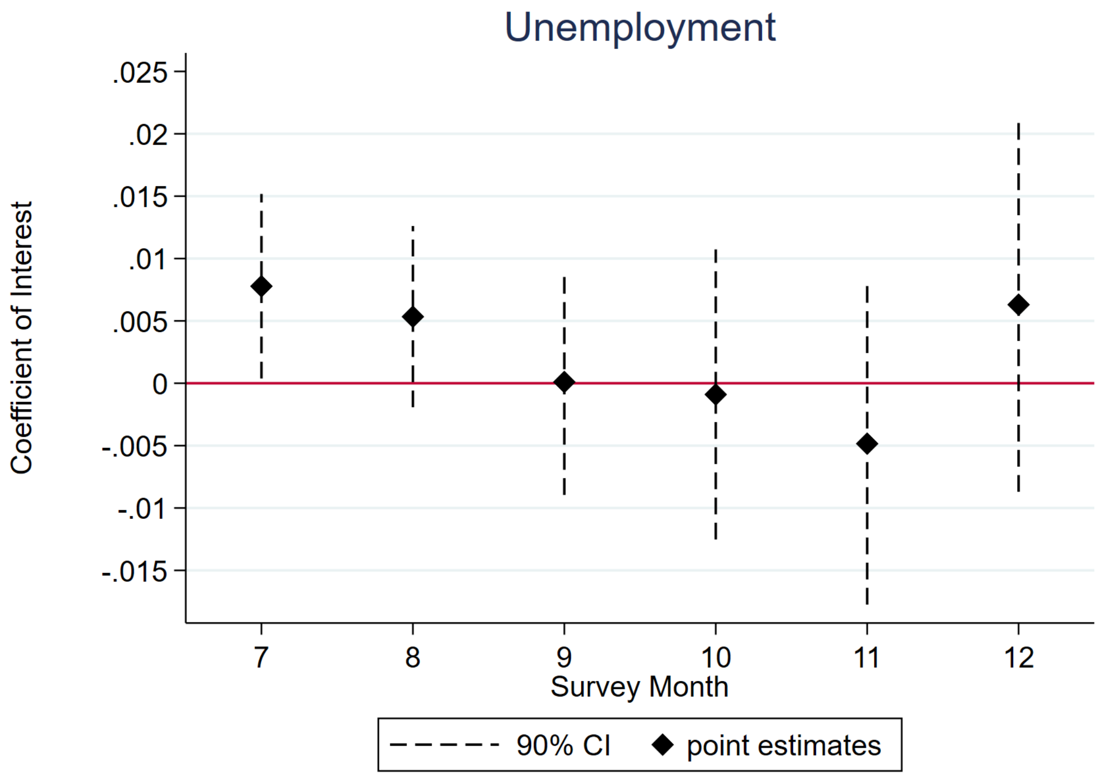
Notes: In this figure we demonstrate the exogeneity of treatment conditional on prefecture characteristics. Blues bars denote the similarity between prefecture A and B. Pink bars denote different timing of case report, which is highly likely driven by random factors. The shadowed area denotes zero-Covid duration.

Figure A5: Conditional-Exogenous Treatment



Notes: The figure shows that conditional on the the number of days with confirmed cases (X-axis), the Duration (Y-axis) is highly likely driven by random factors. Outliers (95 percentile) of Days with New Cases are dropped from graph.

Figure A6: Conditional-Exogenous Treatment 2



*Notes:*The figure shows the estimated effect of zero-covid policy on probability of unemployment, from July to December. Reporting 90% confidence intervals.

Figure A7: Treatment Effect by Survey Month

3.8.3 Appendix Tables

Table A1: Sample by Waves

	B			
	Year	Prefectures	Obs	Share
r1	2014	125	16246	.26
r2	2016	125	17453	.28
r3	2018	123	18379	.29
r4	2020	121	11265	.18
r5	.	.	63343	1

Notes: The table reports the distribution of sample sizes across four waves (2014, 2016, 2018 and 2020).

Bibliography

[1] Daron Acemoglu, Simon Johnson, and James A. Robinson. The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401, December 2001.

[2] Daron Acemoglu and James A Robinson. *The narrow corridor: States, societies, and the fate of liberty*. Penguin, 2020.

[3] Abi Adams-Prassl, Teodora Boneva, Marta Golin, and Christopher Rauh. Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics*, 189:104245, September 2020.

[4] Alberto Alesina and Paola Giuliano. Culture and institutions. *Journal of Economic Literature*, 53(4):898–944, December 2015.

[5] Yann Algan and Pierre Cahuc. Inherited trust and growth. *American Economic Review*, 100(5):2060–92, December 2010.

[6] Annette Alstadsæter, Bernt Bratsberg, Gaute Eielsen, Wojciech Kopczuk, Simen Markussen, Oddbjorn Raaum, and Knut Røed. The First Weeks of the Coronavirus Crisis: Who Got Hit, When and Why? Evidence from Norway. Working Paper 27131, National Bureau of Economic Research, May 2020. Series: Working Paper Series.

[7] Jessica Anania, Bernardo Andretti de Mello, Noam Angrist, Roy Barnes, Thomas Boby, Alice Cavalieri, Benjamin Edwards, Samuel Webster, Lucy Ellen, Rodrigo Furst, Rafael Goldszmidt, Maria Luciano, Saptarshi Majumdar, Radhika Nagesh, Annalena Pott, Andrew Wood, Adam Wade, Hao Zha, Adam Wade, Anna Petherick, Beatriz Kira, Emily Cameron-Blake, Helen Tatlow, Jodie Elms, Kaitlyn Green, Laura Hallas, Martina Di Folco, Thomas Hale, Toby Phillips, and Yuxi Zhang. Variation in government responses to covid-19. https://www.bsg.ox.ac.uk/sites/default/files/2022-04/bsg-wp-2020-032-v13.pdf, December 2022. Blavatnik School Working Paper.

[8] Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Number 8769 in Economics Books. Princeton University Press, 2009.

[9] Noam Angrist, Pinelopi Koujianou Goldberg, and Dean Jolliffe. Why is growth in developing countries so hard to measure? *Journal of Economic Perspectives*, 35(3):215–42, 2021.

[10] Dmitry Arkhangelsky, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. Synthetic Difference-in-Differences. *American Economic Review*, 111(12):4088–4118, December 2021.

[11] Kenneth J Arrow. Gifts and exchanges. *Philosophy & Public Affairs*, page 357, 1972.

[12] Sangmin Aum, Sang Yoon (Tim) Lee, and Yongseok Shin. COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea. *Labour Economics*, 70:101993, June 2021.

[13] Sangmin Aum, Sang Yoon (Tim) Lee, and Yongseok Shin. Inequality of fear and self-quarantine: Is there a trade-off between GDP and public health? *Journal of Public Economics*, 194:104354, February 2021.

[14] Paolo Bajardi, Chiara Poletto, Jose J. Ramasco, Michele Tizzoni, Vittoria Colizza, and Alessandro Vespignani. Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLOS ONE*, 6(1):e16591, January 2011. Publisher: Public Library of Science.

[15] Marianna Belloc, Francesco Drago, and Roberto Galbiati. Earthquakes, Religion, and Transition to Self-Government in Italian Cities*. *The Quarterly Journal of Economics*, 131(4):1875–1926, 07 2016.

[16] John Bellows and Edward Miguel. War and local collective action in sierra leone. *Journal of Public Economics*, 93(11):1144–1157, 2009.

[17] Michaela Benzeval, Jonathan Burton, Thomas F. Crossley, Paul Fisher, Annette Jäckle, Hamish Low, and Brendan Read. The Idiosyncratic Impact of an Aggregate Shock: The Distributional Consequences of COVID-19. SSRN Scholarly Paper ID 3615691, Social Science Research Network, Rochester, NY, May 2020.

[18] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275, 2004.

[19] George J Borjas and Hugh Cassidy. The Adverse Effect of the COVID-19 Labor Market Shock on Immigrant Employment. page 43, 2020.

[20] Abel Brodeur, Nikolai Cook, and Taylor Wright. On the effects of covid-19 safer-at-home policies on social distancing, car crashes and pollution. *Journal of Environmental Economics and Management*, 106:102427, 2021.

[21] Abel Brodeur, David Gray, Anik Islam, and Suraiya Bhuiyan. A literature review of the economics of COVID-19. *Journal of Economic Surveys*, 35(4):1007–1044, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12423.

[22] V. Buchard, A.M. da Silva, C.A. Randles, P. Colarco, R. Ferrare, J. Hair, C. Hostetler, J. Tackett, and D. Winker. Evaluation of the surface pm2.5 in version 1 of the nasa merra aerosol reanalysis over the united states. *Atmospheric Environment*, 125:100–111, 2016.

[23] Johannes C Buggle and Ruben Durante. Climate Risk, Cooperation and the Co-Evolution of Culture and Institutions. *The Economic Journal*, 131(637):1947–1987, 01 2021.

[24] Dzung Bui, Lena Dräger, Bernd Hayo, and Giang Nghiem. The effects of fiscal policy on households during the COVID-19 pandemic: Evidence from Thailand and Vietnam. *World Development*, 153:105828, May 2022.

[25] Louis-Philippe Béland, Abel Brodeur, and Taylor Wright. The Short-Term Economic Consequences of Covid-19: Exposure to Disease, Remote Work and Government Response. SSRN Scholarly Paper ID 3584922, Social Science Research Network, Rochester, NY, April 2020.

[26] Tomaz Cajner, Leland D. Crane, Ryan A. Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz. The U.S. Labor Market during the Beginning of the Pandemic Recession. Working Paper 27159, National Bureau of Economic Research, May 2020. Series: Working Paper Series.

[27] Brantly Callaway, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. Difference-in-Differences with a Continuous Treatment. *arXiv:2107.02637 [econ]*, July 2021. arXiv: 2107.02637.

[28] Filipe Campante and David Yanagizawa-Drott. Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan *. *The Quarterly Journal of Economics*, 130(2):615–658, 01 2015.

[29] Jiarui Cao, Yiqing Xu, and Chuanchuan Zhang. Clans and calamity: How social capital saved lives during china's great famine. *Journal of Development Economics*, 157:102865, 2022.

[30] Vivek Charu, Scott Zeger, Julia Gog, Ottar N. Bjørnstad, Stephen Kissler, Lone Simonsen, Bryan T. Grenfell, and Cécile Viboud. Human mobility and the spatial transmission of influenza in the United States. *PLOS Computational Biology*, 13(2):e1005382, February 2017. Publisher: Public Library of Science.

[31] Haiqiang Chen, Wenlan Qian, and Qiang Wen. The impact of the covid-19 pandemic on consumption: Learning from high-frequency transaction data. *AEA Papers and Proceedings*, 111:307–11, May 2021.

[32] Jiandong Chen, Ming Gao, Shulei Cheng, Wenxuan Hou, Malin Song, Xin Liu, and Yu Liu. Global 1 km× 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *Scientific Data*, 9(1):202, 2022.

[33] Jinjing Chen, Wei Chen, Ernest Liu, Jie Luo, and Zheng Song. The economic cost of locking down like china: Evidence from city-to-city truck flows. https://www.econ.cuhk.edu.hk/econ/images/documents/truck_flow_and_covid19_220315.pdf, 2022. Working Paper.

[34] Shuai Chen, Paulina Oliva, and Peng Zhang. The effect of air pollution on migration: Evidence from china. *Journal of Development Economics*, 156:102833, 2022.

[35] Shuo Chen and Xiaohuan Lan. There will be killing: Collectivization and death of draft animals. *American Economic Journal: Applied Economics*, 9(4):58–77, October 2017.

[36] Yi Chen, Ziying Fan, Xiaomin Gu, and Li-An Zhou. Arrival of young talent: The send-down movement and rural education in china. *American Economic Review*, 110(11):3393–3430, November 2020.

[37] Zhiwu Chen, Zhan Lin, and Xiaoming Zhang. Hedging desperation: How kinship networks reduced cannibalism in historical china. Technical report, working paper, University of Hong Kong, 2022.

[38] Zhiwu. Chen and chicheng Ma. The confucian clan as a risk-sharing institution: How pre-industrial china became the most populous nation. Working paper, SSRN. https://ssrn.com/abstract=3859796, June 2021.

[39] Zhiwu Chen, Chicheng Ma, and Andrew J Sinclair. Banking on the Confucian Clan: Why China Developed Financial Markets so Late. *The Economic Journal*, 132(644):1378–1413, 10 2021.

[40] Raj Chetty, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data. Working Paper 27431, National Bureau of Economic Research, June 2020. Series: Working Paper Series.

[41] Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77, September 2009.

[42] Damian Clarke. Estimating Difference-in-Differences in the Presence of Spillovers. *MPRA Paper*, September 2017. Number: 81604 Publisher: University Library of Munich, Germany.

[43] Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. Labor Markets During the COVID-19 Crisis: A Preliminary View. Working Paper 27017, National Bureau of Economic Research, April 2020. Series: Working Paper Series.

[44] Pierluigi Conzo and Francesco Salustri. A war is forever: The long-run effects of early exposure to world war ii on trust. *European Economic Review*, 120:103313, 2019.

[45] Ben S. Cooper, Richard J. Pitman, W. John Edmunds, and Nigel J. Gay. Delaying the International Spread of Pandemic Influenza. *PLOS Medicine*, 3(6):e212, May 2006. Publisher: Public Library of Science.

[46] Scott Cunningham. Continuous treatment DiD: An application to abortion clinic closures. https://causalinf.substack.com/p/continuous-treatment-did, December 2021.

[47] Hai-Anh H. Dang, Cuong Viet Nguyen, and Calogero Carletto. Did a successful fight against COVID-19 come at a cost? Impacts of the pandemic on employment outcomes in Vietnam. *World Development*, 161:106129, January 2023.

[48] Mark Dincecco and Yuhua Wang. Internal conflict and state development: Evidence from imperial china. *Available at SSRN 3209556*, 2021.

[49] Dave Donaldson and Adam Storeygard. The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–198, 2016.

[50] Christopher D Elvidge, Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu, and Tilottama Ghosh. Viirs night-time lights. *International journal of remote sensing*, 38(21):5860–5879, 2017.

[51] Benjamin Enke. Kinship, Cooperation, and the Evolution of Moral Systems*. *The Quarterly Journal of Economics*, 134(2):953–1019, 01 2019.

[52] Hanming Fang, Chunmian Ge, Hanwei Huang, and Hongbin Li. Pandemics, global supply chains, and local labor demand: Evidence from 100 million posted jobs in china. Working Paper 28072, National Bureau of Economic Research, November 2020.

[53] Hanming Fang, Long Wang, and Yang. Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China. Working Paper 26906, National Bureau of Economic Research, March 2020. Series: Working Paper Series.

[54] Hanming Fang, Long Wang, and Yang Yang. Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in china. *Journal of Public Economics*, 191:104272, 2020.

[55] Xiaotong Fei. *From the soil.* university of California Press, 1992.

[56] Neil M. Ferguson, Derek A. T. Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, July 2006. Number: 7101 Publisher: Nature Publishing Group.

[57] Eliza Forsythe, Lisa B. Kahn, Fabian Lange, and David G. Wiczer. Labor Demand in the time of COVID-19: Evidence from vacancy postings and UI claims. Working Paper 27061, National Bureau of Economic Research, April 2020. Series: Working Paper Series.

[58] Ayman Fouda, Nader Mahmoudi, Naomi Moy, and Francesco Paolucci. The COVID-19 pandemic in Greece, Iceland, New Zealand, and Singapore: Health policies and lessons learned. *Health Policy and Technology*, 9(4):510–524, December 2020.

[59] Shihe Fu, V Brian Viard, and Peng Zhang. Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China. *The Economic Journal*, 131(640):3241–3273, 04 2021.

[60] John Gibson. Better night lights data, for longer. *Oxford Bulletin of Economics and Statistics*, 83(3):770–791, 2021.

[61] Paola Giuliano and Nathan Nunn. Understanding cultural persistence and change. *Review of Economic Studies*, 88(4):1541–1581, 2021 2021.

[62] Paul Goldsmith-Pinkham and Guido W. Imbens. Social Networks and the Identification of Peer Effects. *Journal of Business & Economic Statistics*, 31(3):253–264, 2013. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].

[63] Da Gong, Zhuocheng Shang, Andong Yan, and Qi Zhang. COVID Risk Level Dataset and Zero Covid Policy in China. Working Paper, 2022. https://economics.ucr.edu/wp-content/uploads/2022/11/11-16-2022-DaAndong_Covid-risk-level.pdf.

[64] Da Gong, Andong Yan, and Jialin Yu. Cost of zero-covid: Effects of anti-contagious policy on labor market outcomes in China. Working Paper, Available at SSRN: https://ssrn.com/abstract=4037688 or http://dx.doi.org/10.2139/ssrn.4037688, December 2022.

[65] Andrew Goodman-Bacon and Jan Marcus. Using difference-in-differences to identify causal effects of covid-19 policies. *Survey Research Methods*, 14(2):153–158, Jun. 2020.

[66] Austan Goolsbee and Chad Syverson. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193:104311, January 2021.

[67] Daniel V. Gordon, R. Quentin Grafton, and Stein Ivar Steinshamn. Cross-country effects and policy responses to COVID-19 in 2020: The Nordic countries. *Economic Analysis and Policy*, 71:198–210, September 2021.

[68] Avner Greif and Guido Tabellini. Cultural and institutional bifurcation: China and europe compared. *American Economic Review*, 100(2):135–40, May 2010.

[69] Avner Greif and Guido Tabellini. The clan and the corporation: Sustaining cooperation in china and europe. *Journal of Comparative Economics*, 45(1):1–35, 2017. Institutions and Economic Change.

[70] Luigi Guiso, Paola Sapienza, and Luigi Zingales. Social Capital as Good Culture. *Journal of the European Economic Association*, 6(2-3):295–320, 05 2008.

[71] Luigi Guiso, Paola Sapienza, and Luigi Zingales. Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 124(3):1095–1131, 2009.

[72] Luigi Guiso, Paola Sapienza, and Luigi Zingales. Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436, 2016.

[73] Sumedha Gupta, Laura Montenovo, Thuy Nguyen, Felipe Lozano-Rojas, Ian Schmutte, Kosali Simon, Bruce A. Weinberg, and Coady Wing. Effects of social distancing policy on labor market outcomes. *Contemporary Economic Policy*, 41(1):166–193, 2023.

[74] Thomas Hale, Yuxi Zhang, Hao Zha, Hui Zhou, Wang Lijun, Zhang Zihan, Tan Zijia, and Deng Longmei. Chinese provincial government responses to covid-19. https://www.bsg.ox.ac.uk/sites/default/files/2022-01/bsg-wp-2021-041-v2.pdf, March 2022. Blavatnik School Working Paper.

[75] Z. Hao, J. Jin, R. Xia, S. Tian, W. Yang, Q. Liu, M. Zhu, T. Ma, C. Jing, and Y. Zhang. Ccam: China catchment attributes and meteorology dataset. *Earth System Science Data*, 13(12):5591–5616, 2021.

[76] Mariaflavia Harari. Cities in bad shape: Urban geometry in india. *American Economic Review*, 110(8):2377–2421, 2020.

[77] Guojun He, Yuhang Pan, and Takanao Tanaka. The short-term impacts of COVID-19 lockdown on urban air pollution in China. *Nature Sustainability*, 3(12):1005–1011, December 2020. Number: 12 Publisher: Nature Publishing Group.

[78] J Vernon Henderson, Tim Squires, Adam Storeygard, and David Weil. The global distribution of economic activity: nature, history, and the role of trade. *The Quarterly Journal of Economics*, 133(1):357–406, 2018.

[79] Roland Hodler and Paul A Raschky. Regional favoritism. *The Quarterly Journal of Economics*, 129(2):995–1033, 2014.

[80] T. Déirdre Hollingsworth, Neil M. Ferguson, and Roy M. Anderson. Will travel restrictions control the international spread of pandemic influenza? *Nature Medicine*, 12(5):497–499, May 2006. Number: 5 Publisher: Nature Publishing Group.

[81] Kisho Hoshi, Hiroyuki Kasahara, Ryo Makioka, Michio Suzuki, and Satoshi Tanaka. The heterogeneous effects of COVID-19 on labor markets: People's movement and non-pharmaceutical interventions. *Journal of the Japanese and International Economies*, 63:101170, March 2022.

[82] Solomon Hsiang, Daniel Allen, Sébastien Annan-Phan, Kendon Bell, Ian Bolliger, Trinetta Chong, Hannah Druckenmiller, Luna Yue Huang, Andrew Hultgren, Emma Krasovich, Peiley Lau, Jaecheol Lee, Esther Rolf, Jeanette Tseng, and Tiffany Wu. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584(7820):262–267, August 2020. Number: 7820 Publisher: Nature Publishing Group.

[83] Tao Hu, Weihe Wendy Guan, Xinyan Zhu, Yuanzheng Shao, Lingbo Liu, Jing Du, Hongqiang Liu, Huan Zhou, Jialei Wang, Bing She, et al. Building an open resources repository for covid-19 research. *Data and Information Management*, 4(3):130–147, 2020.

[84] Cheng Huang, Gordon G. Liu, and Zhejin Zhao. Coming out of the pandemic: What have we learned and what should we learn? *China Economic Review*, page 101934, 2023.

[85] Nick Huntington-Klein. *The effect: An introduction to research design and causality*. Chapman and Hall/CRC, 2021.

[86] Thais Lærkholm Jensen and Niels Johannesen. The consumption effects of the 2007–2008 financial crisis: Evidence from households in denmark. *American Economic Review*, 107(11):3386–3414, November 2017.

[87] Hiroyuki Kasahara and Bingjing Li. Grain exports and the causes of china's great famine, 1959–1961: County-level evidence. *Journal of Development Economics*, 146:102513, 2020.

[88] Stephen Knack. Social capital and the quality of government: Evidence from the states. *American Journal of Political Science*, 46(4):772–785, 2002.

[89] Andre Kurmann, Etienne Lalé, and Lien Ta. The Impact of COVID-19 on Small Business Dynamics and Employment: Real-time Estimates With Homebase Data. SSRN Scholarly Paper ID 3896299, Social Science Research Network, Rochester, NY, July 2021.

[90] Shengjie Lai, Nick W. Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R. Floyd, Amy Wesolowski, Mauricio Santillana, Chi Zhang, Xiangjun Du, Hongjie Yu, and Andrew J. Tatem. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*, 585(7825):410–413, September 2020. Number: 7825 Publisher: Nature Publishing Group.

[91] Jeff Larrimore, Jacob Mortenson, and David Splinter. Earnings shocks and stabilization during COVID-19. *Journal of Public Economics*, 206:104597, February 2022.

[92] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S.M. Leung, Eric H.Y. Lau, Jessica Y. Wong, Xuesen Xing, Nijuan Xiang, Yang Wu, Chao Li, Qi Chen, Dan Li, Tian Liu, Jing Zhao, Man Liu, Wenxiao Tu, Chuding Chen, Lianmei Jin, Rui Yang, Qi Wang, Suhua Zhou, Rui Wang, Hui Liu, Yinbo Luo, Yuan Liu, Ge Shao, Huan Li, Zhongfa Tao, Yang Yang, Zhiqiang Deng, Boxi Liu, Zhitao Ma, Yanping Zhang, Guoqing Shi, Tommy T.Y. Lam, Joseph T. Wu, George F. Gao, Benjamin J. Cowling, Bo Yang, Gabriel M. Leung, and Zijian Feng. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, 382(13):1199–1207, March 2020. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2001316.

[93] Xuecao Li, Yuyu Zhou, Min Zhao, and Xia Zhao. A harmonized global nighttime light dataset 1992–2018. *Scientific data*, 7(1):168, 2020.

[94] Shasha Liu, Gaowen Kong, and Dongmin Kong. Effects of the covid-19 on air quality: Human mobility, spillover effects, and city connections. *Environmental and Resource Economics*, 76:635–653, 2020.

[95] Jeremy Mark and Michael Schuman. China's faltering "zero covid" policy: Politics in command, economy in reverse, 2022.

[96] Luis R Martinez. How much should we trust the dictator's gdp growth estimates? *Journal of Political Economy*, 130(10):2731–2769, 2022.

[97] Edouard Mathieu, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser. Coronavirus pandemic (covid-19). *Our World in Data*, 2020. https://ourworldindata.org/coronavirus.

[98] Xin Meng, Nancy Qian, and Pierre Yared. The institutional causes of china's great famine, 1959-61. Working Paper 16361, National Bureau of Economic Research, September 2010.

[99] Xin Meng, Nancy Qian, and Pierre Yared. The Institutional Causes of China's Great Famine, 1959–1961. *The Review of Economic Studies*, 82(4):1568–1611, 04 2015.

[100] Matthew J Menne, Imke Durre, Russell S Vose, Byron E Gleason, and Tamara G Houston. An overview of the global historical climatology network-daily database. *Journal of atmospheric and oceanic technology*, 29(7):897–910, 2012.

[101] Simon Mongey, Laura Pilossoph, and Alexander Weinberg. Which workers bear the burden of social distancing? *The Journal of Economic Inequality*, 19(3):509–526, September 2021.

[102] Jacob Moscona, Nathan Nunn, and James A. Robinson. Keeping it in the family: Lineage organization and the scope of trust in sub-saharan africa. *American Economic Review*, 107(5):565–71, May 2017.

[103] Jacob Moscona, Nathan Nunn, and James A Robinson. Segmentary lineage organization and conflict in sub-saharan africa. *Econometrica*, 88(5):1999–2036, 2020.

[104] Jacob Moscona and Awa Ambra Seck. Age set vs kin: Culture and financial ties in east africa. *Available at SSRN 2856803*, 2021.

[105] Nathan Nunn and Leonard Wantchekon. The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–52, December 2011.

[106] Nathan Nunn and Leonard Wantchekon. The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–52, 2011.

[107] OECD. *OECD Employment Outlook 2020: Worker Security and the COVID-19 Crisis*. Organisation for Economic Co-operation and Development, Paris, 2020.

[108] OECD. *OECD Employment Outlook 2021: Navigating the COVID-19 Crisis and Recovery*. Organisation for Economic Co-operation and Development, Paris, 2021.

[109] Dominic O'Sullivan, Mubarak Rahamathulla, and Manohar Pawar. The Impact and Implications of COVID-19: An Australian Perspective. *The International Journal of Community and Social Development*, 2(2):134–151, June 2020. Publisher: SAGE Publications India.

[110] Kenneth Pomeranz. *The great divergence: China, Europe, and the making of the modern world economy.* Princeton University Press, 2000.

[111] Giacomo A.M. Ponzetto and Ugo Troiano. Social capital, government expenditures and growth. Working Paper 24533, National Bureau of Economic Research, April 2018.

[112] Robert D Putnam. Bowling alone: Americas declining social capital." journal of democracy 6: 65-78. *CITY PUBLICS*, 379, 1995.

[113] Robert D Putnam, Robert Leonardi, and Rafaella Y Nanetti. *Making democracy work: Civic traditions in modern Italy.* Princeton university press, 1993.

[114] Jinlei Qi, Dandan Zhang, Xiang Zhang, Tanakao Takana, Yuhang Pan, Peng Yin, Jiangmei Liu, Shuocen Liu, George F. Gao, Guojun He, and Maigeng Zhou. Short- and medium-term impacts of strict anti-contagion policies on non-COVID-19 mortality in China. *Nature Human Behaviour*, 6(1):55–63, January 2022. Number: 1 Publisher: Nature Publishing Group.

[115] Yun Qiu, Xi Chen, and Wei Shi. Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics*, 33(4):1127–1172, October 2020.

[116] William T Rowe. *Crimson Rain: Seven Centuries of Violence in a Chinese County.* Stanford University Press, 2007.

[117] Lutz Sager and Gregor Singer. Clean identification? The effects of the Clean Air Act on air pollution, exposure disparities and house prices. LSE Research Online Documents on Economics 115528, London School of Economics and Political Science, LSE Library, May 2022.

[118] Jaeung Sim, Daegon Cho, Youngdeok Hwang, and Rahul Telang. Frontiers: virus shook the streaming star: estimating the covid-19 impact on music consumption. *Marketing Science*, 41(1):19–32, 2022.

[119] Samriddhi Singla, Ahmed Eldawy, Tina Diao, Ayan Mukhopadhyay, and Elia Scudiero. The raptor join operator for processing big raster+ vector data. In *Proceedings of the 29th International Conference on Advances in Geographic Information Systems*, pages 324–335, 2021.

[120] Adam Storeygard. Farther on down the road: transport costs, trade and urban growth in sub-saharan africa. *The Review of economic studies*, 83(3):1263–1295, 2016.

[121] Anthony Sudarmawan, Hao Zha, Emily Cameron-Blake, Helen Tatlow, Kaitlyn Green, Martina Di Folco, Thomas Hale, and Toby Phillips. What have we learned from tracking every government policy on COVID-19 for the past two years? *BSG Research Note*, March 2022.

[122] Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, December 2021.

[123] Guido Tabellini. The Scope of Cooperation: Values and Incentives*. *The Quarterly Journal of Economics*, 123(3):905–950, 08 2008.

[124] Guido Tabellini. Culture and institutions: economic development in the regions of europe. *Journal of the European Economic association*, 8(4):677–716, 2010.

[125] T. Talhelm, X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, and S. Kitayama. Large-scale psychological differences within china explained by rice versus wheat agriculture. *Science*, 344(6184):603–608, 2014.

[126] Thomas Talhelm and Alexander S. English. Historically rice-farming societies have tighter social norms in china and worldwide. *Proceedings of the National Academy of Sciences*, 117(33):19816–19824, 2020.

[127] Thomas Talhelm, Xiao Zhang, Shige Oishi, Chen Shimin, Dechao Duan, Xiaoli Lan, and Shinobu Kitayama. Large-scale psychological differences within china explained by rice versus wheat agriculture. *Science*, 344(6184):603–608, 2014.

[128] Huaiyu Tian, Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz U. G. Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, Ben Wang, Peng Yang, Yujun Cui, Yimeng Song, Pai Zheng, Quanyi Wang, Ottar N. Bjornstad, Ruifu Yang, Bryan T. Grenfell, Oliver G. Pybus, and Christopher Dye. An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. *Science*, 368(6491):638–642, 2020.

[129] Katherine A Twohig, Tommy Nyberg, Asad Zaidi, Simon Thelwall, Mary A Sinnathamby, Shirin Aliabadi, Shaun R Seaman, Ross J Harris, Russell Hope, Jamie Lopez-Bernal, et al. Hospital admission and emergency care attendance risk for sars-cov-2 delta (b. 1.617. 2) compared with alpha (b. 1.1. 7) variants of concern: a cohort study. *The Lancet Infectious Diseases*, 22(1):35–42, 2022.

[130] Andrew G. Walder. *Great Leap*, pages 152–179. Harvard University Press, 2015.

[131] Haidong Wang, Katherine R. Paulson, Spencer A. Pease, Stefanie Watson, Haley Comfort, Peng Zheng, Aleksandr Y. Aravkin, Catherine Bisignano, Ryan M. Barber, Tahiya Alam, John E. Fuller, Erin A. May, Darwin Phan Jones, Meghan E. Frisch, Cristiana Abbafati, Christopher Adolph, Adrien Allorant, Joanne O. Amlag, Bree Bang-Jensen, Gregory J. Bertolacci, Sabina S. Bloom, Austin Carter, Emma Castro, Suman Chakrabarti, Jhilik Chattopadhyay, Rebecca M. Cogen, James K. Collins,

Kimberly Cooperrider, Xiaochen Dai, William James Dangel, Farah Daoud, Carolyn Dapper, Amanda Deen, Bruce B. Duncan, Megan Erickson, Samuel B. Ewald, Tatiana Fedosseeva, Alize J. Ferrari, Joseph Jon Frostad, Nancy Fullman, John Gallagher, Amiran Gamkrelidze, Gaorui Guo, Jiawei He, Monika Helak, Nathaniel J. Henry, Erin N. Hulland, Bethany M. Huntley, Maia Kereselidze, Alice Lazzar-Atwood, Kate E. LeGrand, Akiaja Lindstrom, Emily Linebarger, Paulo A. Lotufo, Rafael Lozano, Beatrice Magistro, Deborah Carvalho Malta, Johan Månsson, Ana M. Mantilla Herrera, Fatima Marinho, Alemnesh H. Mirkuzie, Awoke Temesgen Misganaw, Lorenzo Monasta, Paulami Naik, Shuhei Nomura, Edward G. O'Brien, James Kevin O'Halloran, Latera Tesfaye Olana, Samuel M. Ostroff, Louise Penberthy, Robert C. Reiner Jr, Grace Reinke, Antonio Luiz P. Ribeiro, Damian Francesco Santomauro, Maria Inês Schmidt, David H. Shaw, Brittney S. Sheena, Aleksei Sholokhov, Natia Skhvitaridze, Reed J. D. Sorensen, Emma Elizabeth Spurlock, Ruri Syailendrawati, Roman Topor-Madry, Christopher E. Troeger, Rebecca Walcott, Ally Walker, Charles Shey Wiysonge, Nahom Alemseged Worku, Bethany Zigler, David M. Pigott, Mohsen Naghavi, Ali H. Mokdad, Stephen S. Lim, Simon I. Hay, Emmanuela Gakidou, and Christopher J. L. Murray. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *The Lancet*, 399(10334):1513–1536, April 2022. Publisher: Elsevier.

[132] Qi Wang and John E. Taylor. Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster. *PLOS ONE*, 11(1):e0147299, January 2016. Publisher: Public Library of Science.

[133] Yuhua Wang. Comprehensive Catalogue of Chinese Genealogies, 2020.

[134] Melanie Meng Xue. Autocratic rule and social capital: Evidence from imperial china. *Available at SSRN 2856803*, 2021.

[135] Chuanchuan Zhang. Clans, entrepreneurship, and development of the private sector in China. *Journal of Comparative Economics*, 48(1):100–123, 2020.

[136] Dandan Zhang. The impact of lockdown policies on labor market outcomes of the Chinese labor force in 2020: Evidence based on an employee tracking survey. *China Economic Quarterly International*, 1(4):344–360, December 2021.

[137] Qi Zhang and Mingxing Liu. Local political elite, partial reform symptoms, and the business and market environment in rural china. *Business and Politics*, 12(1):1–39, 2010.

[138] Xiaoyu Zhou, Theodore Alysandratos, and Michael Naef. Rice Farming and the Origins of Cooperative Behaviour. *The Economic Journal*, 133(654):2504–2532, 04 2023.

[139] Christian Zimpelmann, Hans-Martin von Gaudecker, Radost Holler, Lena Janys, and Bettina Siflinger. Hours and income dynamics during the Covid-19 pandemic: The case of the Netherlands. *Labour Economics*, 73:102055, December 2021.