Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China[†]

Hanming Fang[‡] Long Wang[§] Yang Yang[¶]

December 31, 2020

[†]We are grateful to the Co-editor (Rebecca Diamond), two anonymous referees, Jerome Adda, Jere Behrman, Xi Chen, Wojciech Kopczuk, Ming Li, Xincheng Qiu, Tsur Somerville, Yao Zeng, Nicolas R. Ziebarth, and virtual seminar participants at Tsinghua University, VoxChina, ABFER-BFI, and HELP! for useful comments and suggestions. All remaining errors are our own.

[‡]Department of Economics, University of Pennsylvania, 133 S. 36th Street, Philadelphia, PA 19104; and the NBER. Email: hanming.fang@econ.upenn.edu.

[§]School of Entrepreneurship and Management, ShanghaiTech University, 393 Middle Huaxia Road, Shanghai, 201210, China. Email: wanglong@shanghaitech.edu.cn.

[¶]CUHK Business School, The Chinese University of Hong Kong, 12 Chak Cheung Street, Hong Kong SAR, China. Email: zoeyang@cuhk.edu.hk.

Human Mobility Restrictions and the Spread of the Novel Coronavirus(2019-nCoV) in China

Abstract

We quantify the causal impact of human mobility restrictions, particularly the lockdown of Wuhan on January 23, 2020, on the containment and delay of the spread of the Novel Coronavirus (2019-nCoV). We employ difference-in-differences (DID) estimations to disentangle the lockdown effect on human mobility reductions from other confounding effects including panic effect, virus effect, and the Spring Festival effect. The lockdown of Wuhan reduced inflows to Wuhan by 76.98%, outflows from Wuhan by 56.31%, and within-Wuhan movements by 55.91%. We also estimate the dynamic effects of up to 22 lagged population inflows from Wuhan and other Hubei cities – the epicenter of the 2019-nCoV outbreak – on the destination cities' new infection cases. We also provide evidence that the enhanced social distancing policies in the 98 Chinese cities outside Hubei province were effective in reducing the impact of the population inflows from the epicenter cities in Hubei province on the spread of 2019-nCoV in the destination cities. We find that in the counterfactual world in which Wuhan were not locked down on January 23, 2020, the COVID-19 cases would be 105.27% higher in the 347 Chinese cities outside Hubei province. Our findings are relevant in the global efforts in pandemic containment.

Keywords: Human Mobility, Lockdown, Social Distancing, 2019-nCoV, COVID-19, Disease Outbreak

JEL Codes: I18, I10.

1 Introduction

Human mobility contributes to the transmission of infectious diseases that pose serious threats to global health. Indeed, many countries restrict human mobility flows as part of their response plans (Bajardi et al., 2011; Wang and Taylor, 2016; Adda, 2016; Charu et al., 2017). However, restrictions on human mobility are controversial not only because of their negative economic impacts, but also because of the uncertainty about their effectiveness in controlling the epidemic. Even if restricting human movement could lead to improvements in disease control and reductions in health risks, it is empirically challenging to quantify the impact of human mobility on the spread of infectious diseases, and to understand the detailed spatial patterns of how the infectious disease spreads. Both granular disease occurrence data and human mobility data (Charu et al., 2017) are hard to obtain; moreover, it is difficult to disentangle the impact of human mobility from other potential contributing factors (Ferguson et al., 2006; Hollingsworth et al., 2006). We exploit the exogenous variations in human mobility created by lockdowns of Chinese cities, and utilize high-quality data sets to study the effectiveness of an unprecedented *cordon sanitaire* of the epicenter of COVID-19, and provide a comprehensive analysis on the role of human mobility restrictions in the delaying and the halting of the spread of the COVID-19 pandemic.[1]

The fast-moving 2019-nCoV that infected 17.5 million people and claimed 682,612 lives as of July 31, 2020 is deteriorating into one of the worst global pandemics.[2] The virus first appeared in Wuhan in early December of 2019, spread mainly through person-to-person contact (Chan et al., 2020), and rapidly reached more than 183 countries as of March 21, 2020.[3]

However, the nature of the virus and its transmission was not publicly known initially. Even though the Chinese state media reported the first known death caused by the Novel Coronavirus on January 11, 2020, it did not register much public attention. On January 14, 2020, the World Health Organization (WHO) announced that Chinese authorities had seen "no clear evidence of human-to-human transmission of the Novel Coronavirus." As the Chinese New Year – which fell on January 25, 2020 – approached, Chinese citizens across the country were still traveling and gathering for various festive and social activities. In fact, on January 18, 2020, more than 10,000 Wuhan families gathered for the annual Wuhan Lunar New Year banquet.[4] These announcements and events suggest that no public panic was sparked yet as of that date. On January 19, 2020, the Chinese National Health Commission convened and sent a team of epidemiologists to Wuhan to investigate the outbreak of the Novel Coronavirus. On January 20, 2020, Dr. Zhong Nanshan, a renowned epidemiologist who was leading the investigative team, for the first time publicly confirmed on

[1]Throughout the paper, we use 2019-nCoV as the official name for the Novel Coronavirus according to the World Health Organization, and use COVID-19 as the name of the disease caused by 2019-nCoV.

[2]Source: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019`

[3]With a population of over 11 million, Wuhan is the largest city in Hubei, the most populous city in Central China, and the seventh most populous city in China. The city is also a major transportation hub, with dozens of railways, roads and expressways passing through and connecting to other major cities, and home to 82 colleges and more than 1 million college students.

[4]See Wei, Lingling, and Chao Deng. "China's Coronavirus Response Is Questioned: 'Everyone Was Blindly Optimistic'." The Wall Street Journal, Dow Jones amp; Company, 24 Jan. 2020, www.wsj.com/articles/china-contends-with-questions-over-response-to-viral-outbreak-11579825832.

national TV that the Novel Coronavirus could spread from human to human. This confirmation caused great panic among the public. The search frequencies in Baidu, the major Chinese search engine, of COVID-19 related keywords such as “Coronavirus”, “Wuhan”, “Bat”, “SARS”, and “symptom” (in Chinese) immediately surged (See Figure A6), and the number of people leaving Wuhan spiked (see Section 3 and Figure A2). On January 23, 2020, the Chinese central government imposed a lockdown in Wuhan, and within a day also on the other cities in Hubei province, in an effort to quarantine the epicenter of the outbreak.

The lockdown of 11 million people in Wuhan represents by then the largest quarantine in public health history, which offers us an opportunity to rigorously examine the effects of the city lockdown and understand the relationship between human mobility and virus transmission. We should point out that the Wuhan lockdown, which lasted 76 days, was extremely strict. All air, train and bus travels into and out of Wuhan were suspended, and all the highway and local road accesses were blocked except for emergency, medical and supply personnel. Residents were not allowed to leave their residence, and food supplies were ordered via phone apps and delivered to the doorsteps by community organizations. The strictness of the lockdown measures was unprecedented.

In this paper, we study four research questions. First, how does the lockdown of Wuhan affect population movement? Second, how do population flows among Chinese cities, particularly outflows from Wuhan, affect the virus infections in the destination cities? Third, are social distancing policies in the destination cities effective in reducing the spread? Fourth, how many infections were prevented elsewhere in China by the unprecedented Wuhan lockdown? We utilize datasets on city-pair population migration and the within-city population movements of each city at the daily level from Baidu Migration, and the city-level daily number of infections from the Chinese Center for Disease Control and Prevention (CCDC) during a sample period of January 1 to February 29, 2020, covering 22 days before and 38 days after the city lockdown on January 23, 2020, as well as the matched data from the same lunar calendar period in 2019.

We first employ difference-in-differences (DID) estimation strategies to disentangle the effect of Wuhan lockdown on human mobility reductions from other confounding effects including panic effect, virus effect, and the Spring Festival effect. Our estimates show that the Wuhan lockdown reduced inflow into Wuhan by 76.98%, outflows from Wuhan by 56.31%, and within-Wuhan movements by 55.91%. We find a clear inverted U-shape relationship between the lagged days of the population inflows from Wuhan or other cities in Hubei and the destination cities' new COVID-19 cases, with the largest impact from the population inflows from the epicenter about 12 to 14 days earlier. We provide evidence that imposing enhanced social distancing policies in 98 Chinese cities outside Hubei Province were effective in reducing the impact of population inflows from the epicenter cities in Hubei province on the spread of 2019-nCoV virus in the destination cities. Finally, we estimate that COVID-19 cases would be 105.27% higher in the 347 Chinese cities outside Hubei province, in the counterfactual world without the Wuhan lockdown.

Our study contributes to a fast-growing literature on 2019-nCoV infection, mostly in the medical and public health fields (Huang et al., 2020; Chan et al., 2020; Liu et al., 2020; WHO, 2003).[5]

[5]This study is also related to disaster-induced migration, which has often occurred during drought, flooding,

Adda (2016) uses a quasi-experimental variation to assess the effectiveness of public health measures that are aimed at reducing interpersonal contacts, and finds that the effectiveness of the measures depends on disease characteristics. Our results are also in line with the results of the latest modeling exercises (Chinazzi et al., 2020; Qiu et al., 2020; Chinazzi et al., 2020; Li et al., 2020; Tian et al., 2020; Jia et al., 2020), which mostly rely on calibrations of various virus-related parameters such as incubation period and detection rates, and changes in travel flows. Lai et al. (2020) build a travel network-based susceptible-exposed-infectious-removed (SEIR) model to simulate the outbreak across cities in mainland China. López and Rodó (2020) also use a modified stochastic SEIR model to explore different post-confinement scenarios in many countries such as Spain, Japan, New Zealand, and the U.S., and they find that the gradual de-confinement could be the best strategy to reduce the disease burden. Recent literature also shows that non-pharmaceutical interventions (NPIs) have a significant effect on reducing the virus reproduction rates in the U.S., U.K, Italy, Spain, France, Australia, New Zealand, and Singapore. The effectiveness of NPIs in reducing the estimated number of infections varies substantially across countries, ranging from only 0.8% in the U.S. to 99.3% in Singapore (Flaxman et al., 2020; Milne and Xie, 2020; Koo et al., 2020; Siedner et al., 2020).

To the best of our knowledge, this paper is the first to provide a *causal* interpretation of the impact of city lockdown on human mobility and the spread of 2019-nCoV, and to clearly disentangle the lockdown effects from other potential contributing factors such as panic and virus effect, as well as the seasonal Spring Festival effect. Although our study focuses exclusively on the impact of human mobility restrictions on the spread of 2019-nCoV virus in China, our estimated results can have general implications for other countries in their fight against the Novel Coronavirus.

The remainder of the paper is structured as follows. In Section 2, we describe the data sets and provide descriptive statistics. In Section 3, we use various DID estimations to evaluate the lockdown effect on population movements. In Section 4, we quantify the impact of lockdown on the national spread of COVID-19. Section 5 concludes.

2 Data and Descriptive Statistics

Population Migration Data. We obtain inter-city population migration data from Baidu Migration, a travel map offered by the largest Chinese search engine, Baidu. The data is based on real-time location records for every smartphone using the company's mapping app, and thus can precisely reflect the population movements between cities.[6] The Baidu Migration data set covers 120,142 pairs of cities per day for 364 Chinese cities between January 12 and March 12 in 2019, and between January 1 and February 29 in 2020. Note that, by the lunar calendar, the data covers the same period of 24 days before and 36 days after the Spring Festivals, respectively for the year 2019 and 2020. The daily inter-city migration data consist of 2,977,899 city-pair observations each

earthquake and other destructive climatic phenomena (Munshi, 2003; Gray and Mueller, 2012; Lu et al., 2012).

[6] It is important to emphasize here that the mobility data is about the movement of people from one city to another based on geo-location services of the smartphones; as such a person flowing out of city A to city B is not necessarily a *resident* of city A, but he/she must have been to city A before moving to city B.

year. In addition, Baidu provides the daily *within-city* mobility data for each city in the sample period, which is a panel consisting of 21,840 city-day level observations each year.

Specifically, the Baidu Migration data provides three migration intensity indicators: the daily *in-migration* index (IMI) of a city, the daily *out-migration* index (OMI) of a city, and the daily *within-city* migration index (WCMI). We convert the three migration indices into the number of population movements using the actual number of inter-city/within-city population flows in Shanghai provided by the National Earth System Science Data Center (NESSDC). Appendix B provides the details of how we convert the Baidu indices into the number of popuplation movements.

Appendix Table A1 presents the summary statistics of the population flows at the city-pair-day level and city-day level. It shows drastic declines in the average inflows, outflows, and within-city migration in 2020, compared to the same lunar period in 2019. The plummeting of the migration statistics due to the Wuhan lockdown is also depicted in Appendix Figure A2.

Outbreak Data. COVID-19 daily case counts are collected from China CDC, which provides daily updates on confirmed, dead, and recovered COVID-19 cases in each city.[7] Forty-one infections were first confirmed in Wuhan on January 10. We plot the geographic distributions of sample cities and cases in Appendix Figure A1. Panel B of Appendix Table A1 presents the summary statistics of COVID-19 data, and Appendix Figure A3 plots the trends of daily statistics of COVID-19 separately for the epicenter city of Wuhan, for other cities in Hubei, and for cities outside of Hubei.[8]

We have many reasons for treating the officially reported COVID-19 cases in Wuhan and other cities of Hubei with caution, and differently from the data of cities outside Hubei. As the epicenter of COVID-19, the health care systems in Wuhan and other cities in Hubei were overwhelmed by the sheer number of patients who needed laboratory testing, especially in the early phases of the virus outbreak. As such, the over-extended medical system in Wuhan and other cities in Hubei might have caused delays in testing the patients who contracted COVID-19; patients who contracted COVID-19 might have been self-recovered or died, before being officially tested; and some who were infected might have been asymptomatic. Also, government officials in the epicenter cities might have incentives to downplay the severity of the outbreak, at least initially. These considerations impact how we use the outbreak data in Section 4.

3 The Impact of Wuhan Lockdown on Population Movements

To suppress the spread of 2019-nCoV, the central government of China imposed an unprecedented lockdown in Wuhan starting from 10 am of January 23, 2020, and in all but one other Hubei cities a day later. As of February 29, 2020, 115 cities in 25 provinces issued different levels

[7]Source: `http://2019nCoV.chinacdc.cn/2019-nCoV/`

[8]The spike of confirmed cases observed on February 12 in Hubei Province is, for the most part, the result of a change in diagnosis classification for which 13,332 clinically (rather than laboratory) confirmed cases were all reported as new cases on February 12, 2020, even though they might have been clinically diagnosed in the preceding weeks. Also on February 13, 2020, a new Communist Party Secretary of Hubei started in his position.

of lockdown policies. Table A2 in the Appendix provides detailed information about the various forms of population mobility control in different cities.

3.1 Empirical Challenges

There are several confounding factors in our attempt to *causally* quantify the impact of lockdown on human mobility, and on the spread of infectious viruses. First, the virus outbreak happens right before the Spring Festival of the Chinese Lunar New Year (which fell on January 25 in 2020), which causes the largest annual human migration every year.[9] Second, the virus itself, even *in the absence* of a mandatory lockdown, may lead to curtailed human movement as people attempt to avoid the exposure to the virus. We refer to this deterrence effect as the *virus effect*. Third, for Wuhan and other nearby cities in Hubei, there is a possible *panic effect,* in reaction to the virus.[10] The panic effect can lead to an increase in the population outflow from the epicenter of the virus outbreak, and a decrease of the population inflow to the epicenter, particularly Wuhan. The panic effect is likely to peak when the government officially confirmed on January 20, 2020 that the Novel Coronavirus can transmit from person-to-person.

3.2 Effects of Various Factors on Inter-City Population Mobility

We first examine the impact of city lockdown on inter-city population mobility, including inflow and outflow, between a city pair (i, j). To disentangle the contributions of these confounding factors on human mobility, we exploit many unique sources of variations in the data, and employ several DID estimation strategies by comparing different treatment and control groups. The DID specification can be described as follows:

$$\begin{aligned} \mathit{Ln(Flow}_{i,j,t}) = &\alpha + \beta_1 \cdot \mathit{Treat} * \mathit{Before}_{1,t} + \beta_2 \cdot \mathit{Treat} * \mathit{Before}_{2,t} \\ &+ \beta_3 \cdot \mathit{Treat} * \mathit{After}_t + \mu_{i,j} + \theta_t + \epsilon_{i,j,t} \end{aligned} \tag{1}$$

where i, j, and t respectively index the destination city, origination city, and date; the dependent variable, $\mathit{Ln(Flow}_{i,j,t})$, is the logarithm of the population flows from city j to city i at date t. The definition of *Treat* varies by specific DID designs, and we will be explicit about its definition below. The city-pair fixed effect $\mu_{i,j}$ is included to absorb the city-specific and the city-pair specific heterogeneities that may contaminate the estimation of our interested coefficient β_3. We also control for the date-fixed effect θ_t to eliminate the time-specific impact, including the Spring Festival travel effect. The standard errors are clustered at the daily level.

In Eq. (1), we include two pre-lockdown period indicators: $\mathit{Before}_{1,t}$ is a dummy that takes value 1 for the period from January 11 to January 19, 2020 (4 to 11 days before the Wuhan lockdown),

[9] The migration across China, which officially begins from about two weeks before, and ends about three weeks after, the Lunar New Year is often referred to as *Chunyun* (meaning Spring movement). In 2019, approximately 3 billion trips were made during Chunyun, see `https://www.cnn.com/travel/article/lunar-new-year-travel-rush-2019/index.html`.

[10] Keane and Neal (2020) find that both virus transmission and government policies can significantly contribute to consumers' panic purchases.

which can be used to examine the parallel pre-trend assumption in the DID analysis; $\mathit{Before}_{2,t}$ is a dummy that takes value 1 for the period from January 20 to January 22, 2020, three days before the unprecedented Wuhan lockdown, but after the official announcement that the virus can spread from person to person. $\mathit{Before}_{2,t}$ allows us to capture the panic effect. Finally, After_t is a dummy that takes value 1 for the sample period after the Wuhan lockdown, between January 23 and February 29, 2020. The omitted benchmark period is from January 1 to January 10, 2020.

3.3 Effects of Various Factors on Within-City Population Mobility

We also estimate the effect of lockdown on the within-city population movement utilizing the city-level data and a variety of DID specifications:

$$\begin{aligned} \mathit{Ln(WithinCityFlow}_{i,t}) =& \alpha + \beta_1 \cdot \mathit{Treat} * \mathit{Before}_{1,t} + \beta_2 \cdot \mathit{Treat} * \mathit{Before}_{2,t} \\ & + \beta_3 \cdot \mathit{Treat} * \mathit{After}_t + \mu_i + \theta_t + \epsilon_{i,t} \end{aligned} \tag{2}$$

where i and t index the city and date. $\mathit{Ln(WithinCityFlow}_{i,t})$ is the logarithm of the within-city population movement for city i at date t. Similar to Eq. (1), *Treat* will be defined according to the DID design. $\mathit{Before}_{1,t}$, $\mathit{Before}_{2,t}$ and After_t are defined in the same way as in Eq. (1). We include the city fixed effects μ_i and date fixed effects θ_t. The standard errors are clustered at the daily level.

3.4 Estimation Results

Table 1 reports the results from three sets of regressions specified according to Eq. (1) for inter-city inflows (Panel A) and outflows (Panel B), and according to Eq. (2) for within-city movement (Panel C). We implement two models that differ in the estimation sample, and the definition of the control group.

[Table 1 About Here]

Model 1: Wuhan 2020 vs. Wuhan 2019. In Model 1, we compare the population movements of Wuhan in 2020 to *itself* in the same matched lunar calendar period in 2019, during which Wuhan is free of virus outbreak and lockdown. Thus, the estimation sample in Model 1 is the daily inflows into and outflows out of Wuhan, as well as the daily within-city movements in Wuhan for years 2019 and 2020.

Under Model 1, *Treat* takes value 1 if the destination city i (respectively, the origination city j) is Wuhan and year is 2020 in Panel A (respectively, Panel B). Panel C examines the within-city mobility, *Treat* takes value 1 if the year is 2020. The control group is Wuhan 2019. We interpret the coefficient estimate of $\mathit{Treat} * \mathit{Before}_{2,t}$ as measuring the panic effect of Wuhan 2020 relative to Wuhan 2019; and the coefficient estimate of $\mathit{Treat} * \mathit{After}_t$ as measuring both the lockdown and the virus effects. The coefficient estimate of $\mathit{Treat} * \mathit{Before}_{1,t}$ examines whether the parallel trend assumption for DID is satisfied. The possibly time-varying Spring Festival effects and the virus effects are both absorbed in the day fixed effects.

The estimated coefficients on $Treat * After_t$ remain negative, and economically and statistically significant in all panels. The estimates suggest that the lockdown of Wuhan, together with the deterrence effect of the virus (the *virus effect*), on average reduces the inflow population into, outflow population from, and within-city movements in Wuhan by 91.94% ($= 1 - exp(-2.518)$), 72.61% ($= 1 - exp(-1.295)$), and 84.45% ($= 1 - exp(-1.861)$), respectively, relative to the same lunar calendar days in 2019. We also find that the coefficient on $Treat * Before_{2,t}$ is significantly positive in Panel B and significantly negative in Panel C, suggesting that the official confirmation of person-to-person spread of COVID-19 creates a panic effect, causing an increase of outflow from Wuhan of 106.06% ($= exp(0.723) - 1$), and a decrease of within-city movements in Wuhan of 24.04% ($= 1 - exp(-0.275)$), during the three days after the confirmation of the person-to-person transmission but before the city lockdown. However, we do not observe a statistically significant panic effect for the population inflow into Wuhan, suggesting that people in other cities were not yet sufficiently concerned about the virus outbreak in Wuhan and did not avoid traveling to Wuhan, even after the official confirmation of the person-to-person transmission. Finally, the coefficient estimates for $Treat * Before_{1,t}$ are all statistically insignificant, suggesting that the parallel pre-trend assumption is plausible.

Model 2: Wuhan 2020 vs. Seven Other Lockdown Cities 2020. In Model 1, the coefficient estimates of $Treat * After_t$ provide an estimate of the *sum* of the lockdown and the virus effects. In order to isolate the lockdown effect from the virus deterrence effect, we consider Model 2, where the estimation sample consists of data of Wuhan and *seven other cities* that went into partial lockdown on February 2 and February 4, 2020, 10 to 12 days after the lockdown of Wuhan.[11] As we show in Table A3 in the Appendix, these seven cities are more comparable to Wuhan than other unlocked cities in terms of the epidemic situation and other economic indicators, and thus provide a more reasonable control group to partial out the virus effect. In particular, it is much more plausible than using cities that were never locked down as control cities to assume that the virus deterrence effect on human mobility in the seven cities is similar to that in Wuhan.

The estimation sample for Model 2 consists of data from Wuhan and the seven cities for the period between January 1 and February 2, 2020. Note that during this period, none of the seven control cities were locked down yet, though they soon would be. The definitions for *Treat* variables are as follows. In Panel A, *Treat* takes value 1 if the destination city i is Wuhan; in Panel B, *Treat* takes value 1 if the origination city j is Wuhan; in Panel C, *Treat* takes value 1 if city i is Wuhan. The control group consists of the seven cities.

We find that the Wuhan lockdown significantly reduces the inflow into, outflow from, and within-city movements in Wuhan by 76.98% ($= 1 - exp(-1.469)$), 56.31% ($= 1 - exp(-0.828)$), and 55.91% ($= 1 - exp(-0.819)$), respectively. We interpret these as the *pure* lockdown effect on population mobility related to Wuhan.

[11] These seven cities are summarized in Appendix Table A2.

Summary. Based on our estimation of Models 1 and 2, Table 2 summarizes our estimates of the panic effect, the virus effect, and the lockdown effect on inflows into, outflows from, and within-city population movements in Wuhan.

[Table 2 About Here]

In Table 2, the lockdown effects are directly calculated from the corresponding coefficient estimates of $Treat * After$ from Model 2 discussed above; the panic effects are from the coefficient estimates of $Treat * Before_2$ in Model 1. For the virus effect, we recognize that the coefficient estimates of $Treat * After$ in Model 1 incorporate both the lockdown and the virus effects. Thus, we calculate the *virus effect* on inflows into Wuhan to be $exp(-2.518 - (-1.469)) - 1 = -64.97\%$, on the outflows from Wuhan to be $exp(-1.295 - (-0.828)) - 1 = -37.31\%$, and on the within-city flow in Wuhan to be $exp(-1.861 - (-0.819)) - 1 = -64.73\%$.

Because our models assume that the different effects enter exponentially in explaining the flows - recall the natural log specifications in Eqs. (1) and (2) - when we would like to calculate the impact of two or more effects on the population flows, we should not simply add the individual effects. For example, the joint impact of the panic and virus effects on outflows out of Wuhan is $(1 + 106.06\%) * (1 - 37.31\%) - 1 = 29.18\%$, instead of the simple sum of $106.06\% + (-37.31\%) = 68.75\%$.

4 Quantify the Impact of Lockdown on the National Spread of COVID-19

4.1 Inter-city Flows and the 2019-nCoV Transmission

We now examine the impact of human mobility on the transmission of 2019-nCoV. Considering that almost all the new COVID-19 cases outside Wuhan were confirmed after the Wuhan lockdown, while almost all inter-city population flows occurred prior to the Wuhan lockdown (see Appendix Figures A2 and A3), we investigate the imported infections by looking specifically at the impact of population inflows from cities in the epicenter, namely, Wuhan and other cities in Hubei province, on the new cases in the destination cities outside Hubei.

Recognizing that 2019-nCoV has a long incubation period, we estimate a *dynamic distributed lag* regression model taking into account that inflows from Wuhan with different lags may have differential impacts on the current new cases in the destination cities. Most of the medical literature states that the 2019-nCoV virus has a median incubation period of five days, and some can have an incubation period of 14 days or more (see, e.g., Lauer et al. (2020)). Luckily, our data allows us to incorporate the possibility that contact with an infected person from Wuhan or other cities in Hubei can result in confirmed infections in the destination city for up to 22 days.

The analysis focuses on the daily new confirmed cases in the post-Wuhan lockdown period from January 23 to February 29, 2020, for cities i that are *outside Hubei province.* Specifically, we run the following regression:[12]

$$
\begin{aligned}
\mathit{Ln(1{+}NewCase}_{i,t}) = \alpha &+ \sum_{\kappa=1}^{22} \beta_{1\kappa} \cdot \mathit{Ln}\left(\mathit{Inflow}_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right) \\
&+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot \mathit{Ln}\left(\sum_{j\neq i, j\neq WH, j\in HB} \mathit{Inflow}_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa} \right) + \mu_i + \theta_t + \epsilon_{it},
\end{aligned}
\tag{3}
$$

where i indexes the cities outside Hubei, and $t \in \{23, ..., 60\}$ indicates the date.[13] $\kappa \in \{1, ..., 22\}$ indicates the time lapsed from the inflows from Wuhan or other Hubei cities until the current date t. $\mathit{Ln(1{+}NewCase}_{i,t})$ is the logarithm of the number of new confirmed cases in city i at date t. $\mathit{Inflow}_{i,WH,t-\kappa}$ and $\sum_{j\neq i, j\neq WH, j\in HB} \mathit{Inflow}_{i,j,t-\kappa}$ are the inflows from Wuhan, and the inflows from the 16 other cities in Hubei to city i, respectively, κ days prior to the focal date t. $I_{WH,t-\kappa}$ and $\sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa}$ respectively represent the *active infected cases* in Wuhan and other cities in Hubei, κ days prior to the focal date t.[14] Therefore, $\mathit{Ln}\left(\mathit{Inflow}_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right)$ measures the impact of the infected fraction of the inflows from Wuhan κ days earlier on the new infections in city i at date t. We control for destination city fixed effects μ_i and date fixed effects θ_t.

Remark 1. *The role of asymptomatic infected individuals in the transmission of 2019-nCOV virus is now well understood (Rothe et al., 2020). Our data, unfortunately, does not contain asymptomatic infection cases. However, to the extent that there is a constant ratio of asymptomatic and symptomatic cases, which seems to be a plausible assumption based on the current literature, our log-log specification using only the number of symptomatic cases would not be affected by not including the asymptomatic cases.*

Note that in this regression we include only cities outside Hubei Province for two reasons. First, Wuhan and other cities in Hubei province are the outbreak epicenter, and we are interested in how population outflows from these cities affect the destination cities' COVID-19 cases. Second, the confirmed COVID-19 cases in Wuhan and other cities in Hubei are likely to be inaccurate due to the reasons aforementioned in Section 2. In contrast, the confirmed cases in other cities are likely to be accurate, as their numbers are not large enough to overwhelm their local health care system; and the incentives to under-report are much weaker in cities outside of Hubei.

The estimated coefficients $\beta_{1\kappa}$ and $\beta_{2\kappa}$ in Eq. (3) respectively represent the impact of the inflows from Wuhan and other cities in Hubei $\kappa \in \{1, ..., 22\}$ days ago on the destination cities' new cases today. They are respectively plotted in the left and right graphs of Panel A of Figure 1. We also fit a spline smoothed curve of the estimated effects of the different lags of inflows from Wuhan and Hubei, which both show a clear inverted U-shape relationship between the lagged days of the

[12]Our log-log specification is based on the classical susceptible-infectious-removed (SIR) model in epidemiology.

[13]Date $t = 23$ indicates the date of January 23, 2020, and $t = 60$ the date of February 29, 2020.

[14]Active infected cases stand for the cumulative active infections (total infections - total healed - total death) in Wuhan at $t - \kappa$. Note that the total populations in Wuhan and other cities in Hubei are constants that are absorbed in α in Eq. (3). Despite the fact that the confirmed cases in cities in Hubei may be under-reported, as we will show in Section 5, the active cases are still informative about the underlying true infection rates.

population inflows from Wuhan or other cities in Hubei and the destination cities' new COVID-19 cases. Interestingly, both graphs show that the largest impact on the newly confirm cases today in Chinese cities outside Hubei comes from the inflow population from Wuhan or other cities in Hubei about 12 to 14 days ago. The pattern exhibited in Figure 1 is consistent with the hypothesis that the incubation period of the 2019-nCoV is up to 12 to 14 days, and also consistent with a shorter incubation period coupled with secondary infections.

[Figure 1 About Here]

4.2 Effect of Social Distancing on Virus Transmission in Destination Cities

Social distancing at the destination cities is crucial in preventing the possibly asymptomatic transmission from the source city (Chinazzi et al., 2020) and quantifying the effect of social distancing on virus transmission is especially relevant in the stage of pre-epidemic community spread. This section studies the effectiveness of social distancing measures in the destination cities in reducing and containing the spread of the virus.

As shown in Appendix Table A2, within weeks after the Wuhan and Hubei lockdowns, various human mobility restrictions were imposed on 98 other Chinese cities outside Hubei. As described in Table A2, the "lockdowns" in destination cities varied in their degree of strictness, from checkpoints at building entrances to establishing quarantine zones, and from public transit shutdowns to strict limits on the inflows into the city, outflows out of the city, as well as within-city population movements. We interpret the human mobility restrictions in the destination cities as an *enhanced social distancing* policy, because the "lockdown" rules in the destination cities are not as strict as those implemented in Wuhan.

Specifically, we estimate the following specification that is a modified version of the regression specification described by Eq. (3):

$$\begin{aligned} \mathit{Ln(1{+}NewCase}_{i,t}) = \alpha &+ \sum_{\kappa=1}^{22} \beta_{1\kappa} \cdot Ln\left(\mathit{Inflow}_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right) \cdot \left(1 - \mathit{Lockdown}_{i,t}\right) \\ &+ \sum_{\kappa=1}^{22} \gamma_{1\kappa} \cdot Ln\left(\mathit{Inflow}_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right) \cdot \mathit{Lockdown}_{i,t} \\ &+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} \mathit{Inflow}_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa}\right) \cdot \left(1 - \mathit{Lockdown}_{i,t}\right) \\ &+ \sum_{\kappa=1}^{22} \gamma_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} \mathit{Inflow}_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa}\right) \cdot \mathit{Lockdown}_{i,t} \\ &+ \mu_i + \theta_t + \epsilon_{it}, \end{aligned} \tag{4}$$

where $\mathit{Lockdown}_{i,t}$ is a dummy that takes value 1 if time t is a date after destination city i's "lockdown" date, if at all; and 0, otherwise, where the lockdown dates of the 98 cities outside

Hubei are listed in Table A2. If city i never implemented any formal lockdown policy, the dummy is always 0. Therefore, the coefficients, $\beta_{1\kappa}$ and $\beta_{2\kappa}$, respectively, measure the impact of the lagged inflows from Wuhan and other cities in Hubei κ days earlier on the destination cities' current new cases before the city's imposition of its "lockdown," while $\gamma_{1\kappa}$ and $\gamma_{2\kappa}$ represent the effect on destination cities' current new cases of the inflows from Wuhan and other Hubei cities after the imposition of the city's lockdown. If enhanced social distancing that comes from the "lockdown policies" imposed at the 98 Chinese cities outside Hubei is effective in reducing the spread of the virus from population flows from the epicenter of the virus, then we expect that $\gamma_{1\kappa}$ and $\gamma_{2\kappa}$ to be smaller than $\beta_{1\kappa}$ and $\beta_{2\kappa}$, respectively.

We plot the estimated coefficients $\beta_{1\kappa}$ and $\gamma_{1\kappa}$ (respectively, $\beta_{2\kappa}$ and $\gamma_{2\kappa}$) in Panel A (Panel B) of Appendix Figure A4, and their differences in Panel B of Figure 1 for the lagged effects of inflows from Wuhan (respectively, 16 non-Wuhan cities in Hubei) on the daily new cases in destination cities outside Hubei. The estimated lagged effects of inflows from Wuhan and other cities in Hubei before the destination city's lockdown, if any, show little change compared to the coefficients in Panel A of Figure 1; however, the coefficient estimates of lagged inflows after the destination city's lockdown policies appear to be statistically insignificant on almost all lags. As shown in Panel B of Figure 1, the differences between the estimated effects pre and post the destination cities' lockdown policies are positive and statistically significant at 10% or lower level for eight (respectively, three) of the first ten lagged population inflows from Wuhan (respectively, 16 other cities of Hubei).

These results suggest that the enhanced social distancing policies in the destination cities are effective in reducing the impact of population inflows from the source cities of Wuhan and other cities in Hubei on the spread of 2019-nCoV virus in the destination cities. This in turn implies that population inflows from the epicenter contribute to the spread of infection in the destination cities only *before* the social distancing measures are applied; it appears that after implementing their various control measures, cities adopting an enhanced social distancing policies can flatten the upward trajectory of the virus infections.

4.3 How Many COVID-19 Cases Were Prevented by the Wuhan Lockdown?

We next estimate the *counterfactual* number of COVID-19 cases that would have occurred in other cities in the absence of Wuhan lockdown, which would, in turn, require a counterfactual estimate of the outflows from Wuhan to other Chinese cities, had there been no lockdown of Wuhan. As shown in Table 2, in the absence of the Wuhan lockdown, the virus effect and panic effects would have led to a 37.31% decrease and a 106.06% increase in the outflow population from Wuhan, respectively. Therefore, we would expect the outflows from Wuhan after January 23 to be

$$\overbrace{(1-0.3731)}^{\text{Reducton from Virus Effect}} \quad * \quad \overbrace{(1+1.0606)}^{\text{Increase from Panic Effect}} \quad = 1.29 \tag{5}$$

times the *normal outflows* from Wuhan to other cities in the counterfactual world.

We use the daily level of outflows from Wuhan to a city *in 2019 on the same lunar calendar* *day* as a measure of the normal outflow, and multiple the number by 1.29 to obtain the daily counterfactual inflows from Wuhan to the city, had there been no lockdown of Wuhan from January 23, 2020. Using this estimation method, we find that on average, the estimated counterfactual outflows from Wuhan to the 347 cities outside Hubei between January 23 and February 29, 2020 would be 2,848,496, 5.99 times the actual population outflow to those cities during the same period, which is 475,542.[15]

We denote the counterfactual inflows from $j = Wuhan$ into city i at date $s \in \{23, ..., 60\}$ from the above calculation as $\widetilde{Inflow}_{i,WH,s}$. We assume that the Wuhan lockdown did not affect the within-city and inter-city population movements in other cities, as well as the infected cases in Wuhan. We also assume that all the control measures implemented by other cities after the Wuhan lockdown remain in place. Thus the parameter estimates of the dynamic lag effects of inflows from Wuhan and other cities in Hubei, estimated in Eq. (3), remain valid as an epidemiological diffusion equation that is not affected by human mobility restrictions that result from the Wuhan lockdown; the lockdown only affected the human flows.

With these considerations in mind, we simulate the counterfactual number of COVID-19 cases, had there been no Wuhan lockdown, on date $t \in \{23, ..., 60\}$ (i.e., from January 23 to February 29, 2020) in cities i outside Hubei by the following equation:

$$
\begin{aligned}
Ln(1 + \widetilde{NewCase}_{i,t}) = \hat{\alpha} &+ \sum_{\kappa=1}^{22} \hat{\beta}_{1\kappa} \cdot Ln\left(\widetilde{Inflow}_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right) \\
&+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa}\right) + \hat{\mu_i} + \hat{\theta}_t + \hat{\epsilon_{i,t}},
\end{aligned}
\tag{6}
$$

where $\hat{\beta}_{1\kappa}$ and $\hat{\beta}_{2\kappa}$ are coefficient estimates obtained from regressions specified in Eq. (3) and $\hat{\mu}_i$ and $\hat{\theta}_t$ are the estimated city fixed effects and date fixed effects, respectively, from the same regression. Note that in predicting the counterfactual infections without the Wuhan lockdown, we use the *counterfactual* inflows from Wuhan to city i for days after January 23 $\widetilde{Inflow}_{i,WH,s}$ discussed previously.

In Figure 2, we present the counterfactual estimates (in the dotted curve) of COVID-19 cases had there been no Wuhan lockdown, and the officially reported cases (in the solid curve) for cities outside Hubei. The gap between the estimated counterfactual number of infection cases and the officially reported cases on the bottom figure represents the number of COVID-19 cases prevented by the Wuhan lockdown. As of February 29, 2020, the officially reported number of COVID-19 in the 347 cities outside Hubei province was 12,623, but our counterfactual simulation suggests that there would have been around 25,911 cases, had there been no Wuhan lockdown. That is, the COVID-19 cases would be 105.27% higher in 347 cities outside Hubei as of February 29, in the counterfactual world in which the city of Wuhan were not locked down from January 23, 2020.

[15] Recall that outflows from Wuhan are not just residents of Wuhan; any travelers who entered Wuhan for whatever reason and then leave Wuhan would be included in the Wuhan outflows measured by Baidu Migration data.

Our findings thus suggest that the lockdown of the city of Wuhan from January 23, 2020 played a crucial role in reducing the imported infections in other Chinese cities and halted the spread of 2019-nCoV virus.

We show that, the daily new infections would gradually decline and stabilize at around 400 cases from February 18 onwards in the counterfactual world. In the short term, the estimated daily new cases in the counterfactual world would not converge to the real world reported cases. The results suggest that the social distancing measures implemented elsewhere in China would not work as well in containing the spread of 2019-nCOV virus unless Wuhan was locked down on January 23, 2020. It also reflects the challenging situations many countries experienced in containing the spread of virus, where strict lockdowns of the virus epicenters were not imposed early.

[Figure 2 About Here]

We also attempt to estimate the magnitude of the undocumented infection cases in Wuhan and other cities in Hubei province during the early stages of the epidemic. In Appendix C we present the methodology and results. We find that there were substantial undocumented infection cases in the early days of the 2019-nCoV outbreak in Wuhan and other cities of Hubei province, but over time, the gap between the officially reported cases and our estimated "actual" cases narrowed significantly.

5 Conclusion

In this paper, we provide valuable causal evidence on the role of human mobility restrictions on the containment and delay of the spread of contagious viruses, including the 2019-nCoV virus that is now ravaging the world. We find that the lockdown of Wuhan reduced inflows to Wuhan by 76.98%, outflows from Wuhan by 56.31%, and within-Wuhan movements by 55.91%; we also find a substantial virus deterrence effect on population mobility. We estimate the dynamic effects of up to 22 lagged population inflows from Wuhan and other Hubei cities, the epicenter of the 2019-nCoV outbreak, on the destination cities' new infection cases. We find that the lockdown of Wuhan significantly reduced the population mobility and that the enhanced social distancing policies in the destination cities effectively reduce the impact of population inflows from the epicenter cities in Hubei on the spread of 2019-nCoV. Using counterfactual simulations with these estimates, we find that the lockdown contributed significantly to reducing the total infections outside Hubei. Specifically, we find that in the counterfactual world in which the city of Wuhan were not locked down from January 23, 2020, infections would be 105.27% higher in the 347 Chinese cities outside Hubei province. Lockdowns of virus epicenters, if they can be identified before the virus widely spreads, may be an important policy instrument in the fight to contain a fast-moving pandemic.

References

Adda, J. (2016). Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data. *The Quarterly Journal of Economics 131*(2), 891–941.

Bajardi, P., C. Poletto, J. J. Ramasco, M. Tizzoni, V. Colizza, and A. Vespignani (2011). Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PloS one 6*(1).

Chan, J. F.-W., S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon, et al. (2020). A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-person Transmission: a Study of a Family Cluster. *The Lancet 395*(10223), 514–523.

Charu, V., S. Zeger, J. Gog, O. N. Bjørnstad, S. Kissler, L. Simonsen, B. T. Grenfell, and C. Viboud (2017). Human Mobility and the Spatial Transmission of Influenza in the United States. *PLoS computational biology 13*(2), e1005382.

Chinazzi, M., J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, et al. (2020). The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak. *Science*.

Ferguson, N. M., D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke (2006). Strategies for Mitigating an Influenza Pandemic. *Nature 442*(7101), 448–452.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, et al. (2020). Estimating the Effects of Non-Pharmaceutical Interventions on COVID-19 in Europe. *Nature*, 1–5.

Gray, C. L. and V. Mueller (2012). Natural Disasters and Population Mobility in Banglades. *Proceedings of the National Academy of Sciences 109*(16), 6000–6005.

Hollingsworth, T. D., N. M. Ferguson, and R. M. Anderson (2006). Will Travel Restrictions Control the International Spread of Pandemic Influenza? *Nature medicine 12*(5), 497–499.

Huang, C., Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al. (2020). Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *The Lancet 395*(10223), 497–506.

Jia, J. S., X. Lu, Y. Yuan, G. Xu, J. Jia, and N. A. Christakis (2020). Population Flow Drives Spatio-temporal Distribution of COVID-19 in China. *Nature*, 1–5.

Keane, M. P. and T. Neal (2020). Consumer Panic in the COVID-19 Pandemic. *SSRN Working Paper*.

Koo, J. R., A. R. Cook, M. Park, Y. Sun, H. Sun, J. T. Lim, C. Tam, and B. L. Dickens (2020). Interventions to Mitigate Early Spread of SARS-CoV-2 in Singapore: A Modelling Study. *The Lancet Infectious Diseases*.

Lai, S., N. W. Ruktanonchai, L. Zhou, O. Prosper, W. Luo, J. R. Floyd, A. Wesolowski, M. Santillana, C. Zhang, X. Du, H. Yu, and T. A. J (2020). Effect of Non-pharmaceutical Interventions for Containing the COVID-19 Outbreak in China. *doi: https://doi.org/10.1101/2020.03.03.20029843, medRxiv*.

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) from Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*.

Li, R., S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman (2020). Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV2). *Science*.

Liu, Y., A. A. Gayle, A. Wilder-Smith, and J. Rocklöv (2020). The Reproductive Number of COVID-19 is Higher Compared to SARS Coronavirus. *Journal of Travel Medicine*.

López, L. and X. Rodó (2020). The End of Social Confinement and COVID-19 Re-emergence Risk. *Nature Human Behaviour 4*(7), 746–755.

Lu, X., L. Bengtsson, and P. Holme (2012). Predictability of Population Displacement after the 2010 Haiti Earthquake. *Proceedings of the National Academy of Sciences 109*(29), 11576–11581.

Milne, G. J. and S. Xie (2020). The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: A Modelling Analysis. *MedRxiv*.

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the US Labor Market. *The Quarterly Journal of Economics 118*(2), 549–599.

Qiu, Y., X. Chen, and W. Shi (2020). Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China. *Journal of Population Economics*, 1.

Rothe, C., M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, et al. (2020). Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. *New England Journal of Medicine*.

Siedner, M. J., G. Harling, Z. Reynolds, R. F. Gilbert, A. Venkataramani, and A. C. Tsai (2020). Social Distancing to Slow the US COVID-19 Epidemic: An Interrupted Time-series Analysis. *MedRxiv*.

Tian, H., Y. Liu, Y. Li, C.-H. Wu, B. Chen, M. U. Kraemer, B. Li, J. Cai, B. Xu, Q. Yang, et al. (2020). An Investigation of Transmission Control Measures during the First 50 days of the COVID-19 Epidemic in China. *Science 368*(6491), 638–642.

Wang, Q. and J. E. Taylor (2016). Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster. *PLoS one 11*(1).

WHO (2003). Consensus Document on the Epidemiology of Severe Acute Respiratory Syndrome (SARS). Technical report, World Health Organization.

Table 1: The Effects of Wuhan Lockdown on Population Movement

Treatment Group	Wuhan 2020	Wuhan 2020
Control Group	Wuhan 2019	7 Cities 2020
Treatment Effects	Lockdown + Virus	Lockdown
Model	(1)	(2)
Panel A. Dep. Variable: *ln(Inflow Population)*		
Treat*Before$_1$	-0.138	0.149
	(0.098)	(0.151)
Treat*Before$_2$	-0.122	0.239
	(0.083)	(0.153)
Treat*After	-2.518***	-1.469***
	(0.212)	(0.338)
Observations	26,494	72,094
R-squared	0.803	0.761
Panel B. Dep. Variable: *ln(Outflow Population)*		
Treat*Before$_1$	0.042	0.010
	(0.079)	(0.071)
Treat*Before$_2$	0.723***	0.174**
	(0.071)	(0.071)
Treat*After	-1.295***	-0.828***
	(0.155)	(0.168)
Observations	26,326	71,533
R-squared	0.899	0.901
Panel C. Dep. Variable: *ln(Within-city Population Flow)*		
Treat*Before$_1$	-0.014	-0.015
	(0.065)	(0.014)
Treat*Before$_2$	-0.275***	-0.096
	(0.067)	(0.057)
Treat*After	-1.861***	-0.819***
	(0.093)	(0.047)
Observations	120	256
R-squared	0.952	0.938
Fixed Effects	City-Pair FE and Daily FE in Panel A and Panel B City FE and Daily FE in Panel C	

Notes: This table reports the results of estimating Equations (1) and (2). The control and treatment groups for Models 1-2 are described in the text. Fixed effects of city-pair and daily are included in all columns in Panels A and B, and fixed effects of city and daily are included in all columns in Panel C. standard errors are clustered at the daily level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2: Summarizing the Panic Effect, Virus Effect and Lockdown Effect

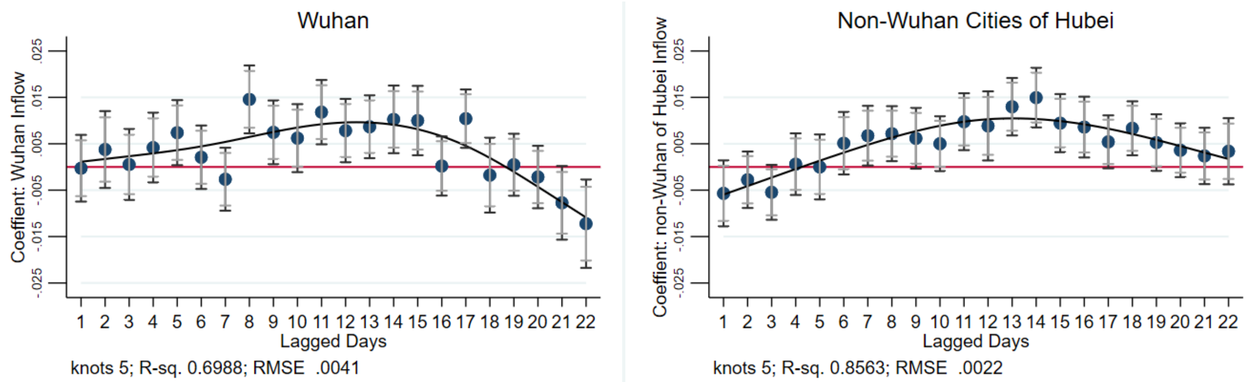
on Inter-City and Within-City Population Movements of Wuhan

Effect	Infows	Outflows	Within-City
Panic Effect	-11.49%	106.06%***	-24.04%***
Virus Effect	-64.97%***	-37.31%***	-64.73%***
Lockdown Effect	-76.98%***	-56.31%***	-55.91%***

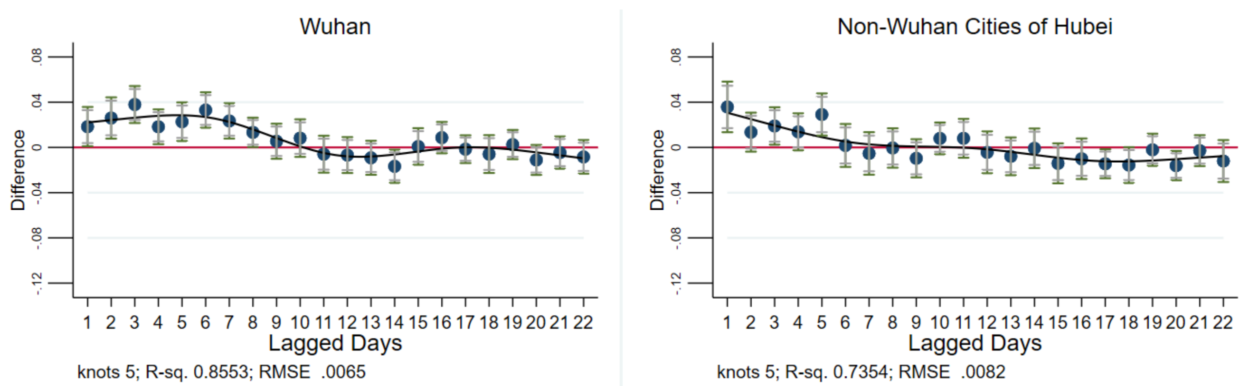
Notes: These effects are calculated based on the estimates reported in Columns (1) and (2) of Table 1. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Figure 1: Dynamic Impact of Past Inflow from Wuhan and from Other Cities in Hubei on Daily New Cases

Panel A: Results of Estimating Equation (3)



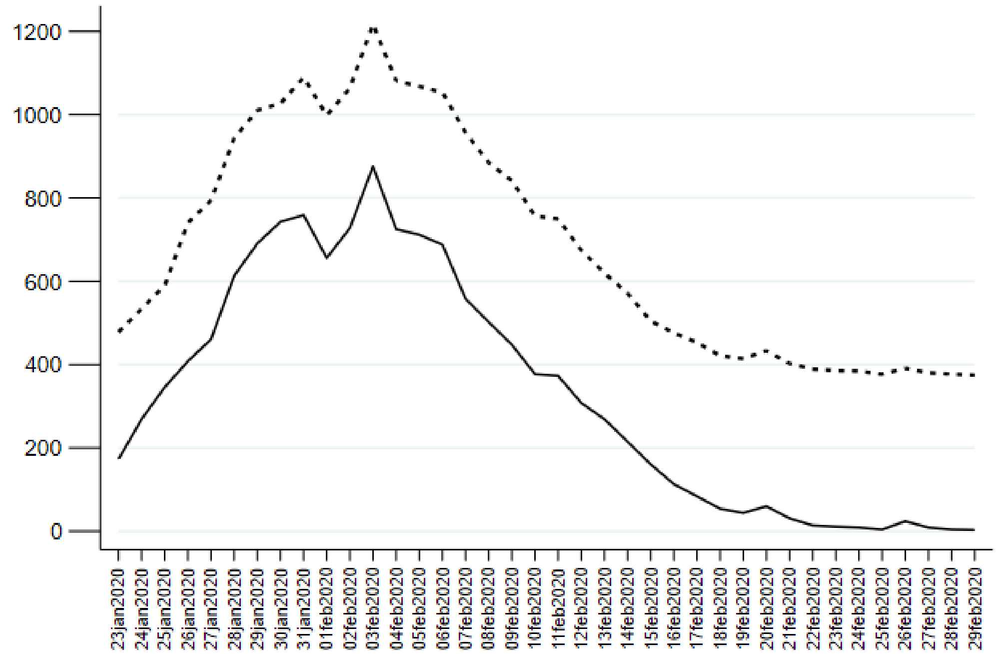
Panel B: Results of Estimating Equation (4)



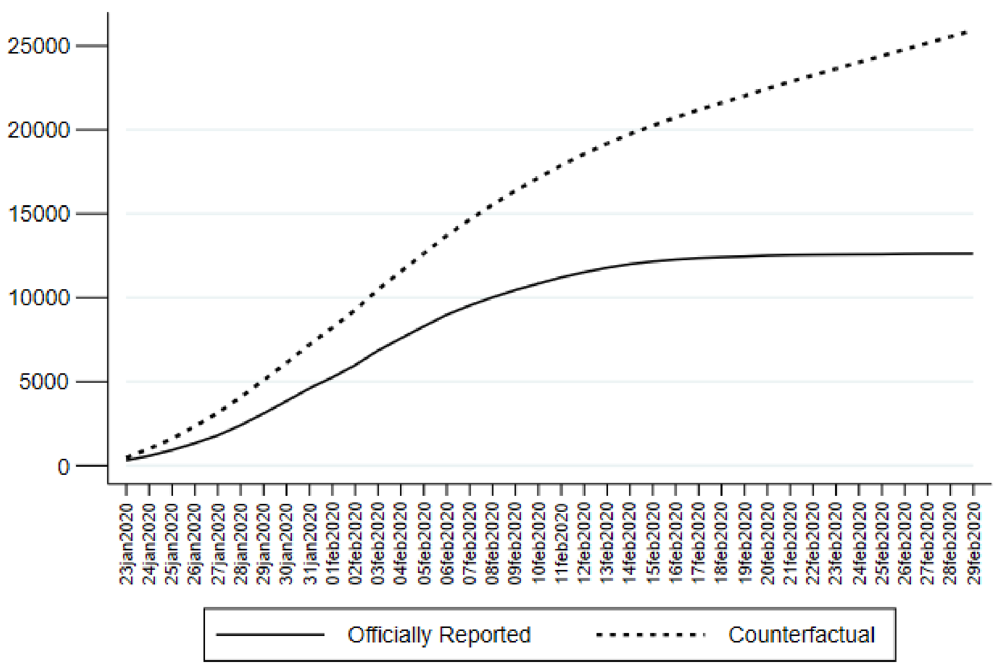
Notes: Panel A plots the dynamic effects of lagged inflows from Wuhan (left) and 16 other cities in Hubei (right) from estimating Equation (3). Panel B plots the difference between the estimated effects of pre- and post-destination-lockdown inflows from Wuhan (left figure) and non-Wuhan cities in Hubei (right figure) on daily new cases in destination cities outside Hubei from estimating Equation (4). We add spline smoothing fit curves (in red) using the *rcspline* function and plot the 90% (the vertical gray whiskers) and 95% (the vertical black whiskers) confidence intervals.

Figure 2: Counterfactual of Infected Cases Elsewhere in China

Panel A: Daily Cases



Panel B: Cumulative Cases



Notes: This figure plots the counterfactual estimation on the COVID-19 cases in 347 other cities in China if Wuhan had never been under a government-ordered lockdown (in the dotted curve), and traces the officially reported COVID-19 cases in cities outside Hubei (in the solid curve). The top figure depicts the model's counterfactual prediction and the actual of daily infection cases, and the bottom figure depicts the evolution of cumulative cases from January 23 to February 29, 2020.

Appendix A. Supplementary Figures and Tables

Table A1. Descriptive Statistics

Panel A: City-Pair-daily Level Data

Sample Period	Jan 12, 2019 - Mar 12 , 2019		Jan 1, 2020 - Feb 29, 2020	
	Mean	S.D	Mean	S.D
CityPair Intensity	0.011	0.052	0.007	0.044
CityPair Flow Population	997.044	4,722.465	668.792	3,970.201

Panel B: City-daily Level Data

Sample Period	Jan 12, 2019 - Mar 12 , 2019		Jan 1, 2020 - Feb 29, 2020	
	Mean	S.D	Mean	S.D
Outflow Intensity	1.159	1.568	0.717	1.478
Outflow Population	105,287.503	142,443.319	65,126.768	134,310.998
Inflow Intensity	1.159	1.608	0.717	1.058
Inflow Population	105,261.037	146,127.871	65,142.334	96,153.964
WithinCity Flow Intensity	4.472	0.818	3.592	1.493
WithinCity Flow Population	1,492,950.639	1,346,858.619	1,188,276.693	1,223,892.324
# of New Confirmed Cases	0	0	3.721	105.176
# of Total Confirmed Cases	0	0	78.364	1,313.831
# of New Deaths	0	0	0.129	2.960
# of Total Deaths	0	0	2.181	50.409
# of New Heals	0	0	1.792	31.757
# of Total Heals	0	0	17.211	259.659

Notes: This table presents the descriptive statistics of the variables used in this study, with Panel A corresponding to the city-pair daily level variables and Panel B corresponding to the city daily level variables. The sample covers a period between Jan 12 and Mar 12 in 2019, and between Jan 1 and Feb 29 in 2020. The two periods in 2019 and 2020 cover the same lunar calendar.

Table A2. Various Levels of Prevention and Control Measures in Different Cities

City	Province	Start Date	Cases as of Feb 29, 2020
Panel A. Complete Shutdown			
Wuhan	Hubei	2020/1/23	49122
Huanggang	Hubei	2020/1/23	2905
Ezhou	Hubei	2020/1/23	1391
Xiaogan	Hubei	2020/1/24	3518
Jingzhou	Hubei	2020/1/24	1579
Suizhou	Hubei	2020/1/24	1307
Huangshi	Hubei	2020/1/24	1014
Yichang	Hubei	2020/1/24	931
Jingmen	Hubei	2020/1/24	925
Xianning	Hubei	2020/1/24	836
Shiyan	Hubei	2020/1/24	672
Xiantao	Hubei	2020/1/24	575
Tianmen	Hubei	2020/1/24	496
Enshi	Hubei	2020/1/24	252
Qianjiang	Hubei	2020/1/24	198
Shennongjia	Hubei	2020/1/24	11
Xiangyang	Hubei	2020/1/28	1175
Panel B. Partital Shutdown			
Wenzhou	Zhejiang	2020/2/2	504
Haerbin	Heilongjiang	2020/2/4	198
Hangzhou	Zhejiang	2020/2/4	169
Ningbo	Zhejiang	2020/2/4	157
Zhengzhou	Henan	2020/2/4	157
Zhumadian	Henan	2020/2/4	139
Fuzhou	Fujian	2020/2/4	72
Panel C. Checkpoints and Quarantine Zones			
Chongqing	Chongqing	2020/1/31	576
Yinchuan	Ningxia	2020/1/31	35
Wuzhong	Ningxia	2020/1/31	28
Fangchenggang	Guangxi	2020/2/2	19
Huaian	Jiangsu	2020/2/3	66
Huaibei	Anhui	2020/2/3	27
Xinyang	Henan	2020/2/4	274
Nanjing	Jiangsu	2020/2/4	93
Xuzhou	Jiangsu	2020/2/4	79
Changzhou	Jiangsu	2020/2/4	51
Linyi	Shandong	2020/2/4	49
Nantong	Jiangsu	2020/2/4	40
Zhenjiang	Jiangsu	2020/2/4	12
Jingdezhen	Jiangxi	2020/2/4	6
Jining	Shandong	2020/2/5	258
Nanchang	Jiangxi	2020/2/5	231
Qingdao	Shandong	2020/2/5	60
Nanning	Guangxi	2020/2/5	55
Sanya	Hainan	2020/2/5	54
Kunming	Yunnan	2020/2/5	53
Cangzhou	Hebei	2020/2/5	48
Jinan	Shandong	2020/2/5	47
Haikou	Hainan	2020/2/5	39
Taizhou	Jiangsu	2020/2/5	37
Taian	Shandong	2020/2/5	35
Shijiazhuang	Hebei	2020/2/5	29
Zaozhuang	Shandong	2020/2/5	24
Yangzhou	Jiangsu	2020/2/5	23
Rizhao	Shandong	2020/2/5	16
Suqian	Jiangsu	2020/2/5	13
Dongying	Shandong	2020/2/5	0
Xinyu	Jiangxi	2020/2/6	130
Shangrao	Jiangxi	2020/2/6	123
Panel C. Checkpoints and Quarantine Zones			
Jiujiang	Jiangxi	2020/2/6	118
Yichun	Jiangxi	2020/2/6	106
Zhuhai	Guangdong	2020/2/6	98
Suzhou	Jiangsu	2020/2/6	87
Ganzhou	Jiangxi	2020/2/6	76
Maanshan	Anhui	2020/2/6	38
Pingxiang	Jiangxi	2020/2/6	33
Shenyang	Liaoning	2020/2/6	28
Jian	Jiangxi	2020/2/6	22
Neijiang	Sichuan	2020/2/6	22
Dalian	Liaoning	2020/2/6	19
Yingtan	Jiangxi	2020/2/6	18
Jinzhou	Liaoning	2020/2/6	12
Yibin	Sichuan	2020/2/6	12
Huludao	Liaoning	2020/2/6	12
Panjin	Liaoning	2020/2/6	11
Dandong	Liaoning	2020/2/6	8
Yaan	Sichuan	2020/2/6	7
Tieling	Liaoning	2020/2/6	7
Chaoyang	Liaoning	2020/2/6	6
Anshun	Guizhou	2020/2/6	4
Liaoyang	Liaoning	2020/2/6	3
Benxi	Liaoning	2020/2/6	3
Fushun	Liaoning	2020/2/6	0
Shenzhen	Guangdong	2020/2/7	417
Guangzhou	Guangdong	2020/2/7	346
Hefei	Anhui	2020/2/7	174
Chengdu	Sichuan	2020/2/7	143
Tianjin	Tianjin	2020/2/7	136
Tangshan	Hebei	2020/2/7	58
Lianyungang	Jiangsu	2020/2/7	48
Lanzhou	Gansu	2020/2/7	36
Guiyang	Guizhou	2020/2/7	36
Suining	Sichuan	2020/2/7	17
Guangyuan	Sichuan	2020/2/7	6
Foshan	Guangdong	2020/2/8	84
Qinhuangdao	Hebei	2020/2/8	10
Ziyang	Sichuan	2020/2/8	4
Dongguan	Guangdong	2020/2/9	99
Huizhou	Guangdong	2020/2/9	62
Wuxi	Jiangsu	2020/2/9	55
Hanzhong	Sanxi	2020/2/9	26
Mianyang	Sichuan	2020/2/9	22
Deyang	Sichuan	2020/2/9	18
Beijing	Beijing	2020/2/10	413
Shanghai	Shanghai	2020/2/10	337
Baotou	Inner Mongolia	2020/2/12	11
Ereduosi	Inner Mongolia	2020/2/12	11
Xilinguole	Inner Mongolia	2020/2/12	9
Chifeng	Inner Mongolia	2020/2/12	9
Bayannaoer	Inner Mongolia	2020/2/12	8
Hulunbeier	Inner Mongolia	2020/2/12	7
Huhehaote	Inner Mongolia	2020/2/12	7
Tongliao	Inner Mongolia	2020/2/12	7
Wulanchabu	Inner Mongolia	2020/2/12	3
Wuhai	Inner Mongolia	2020/2/12	2
Xingan	Inner Mongolia	2020/2/12	1
Alashan	Inner Mongolia	2020/2/12	0

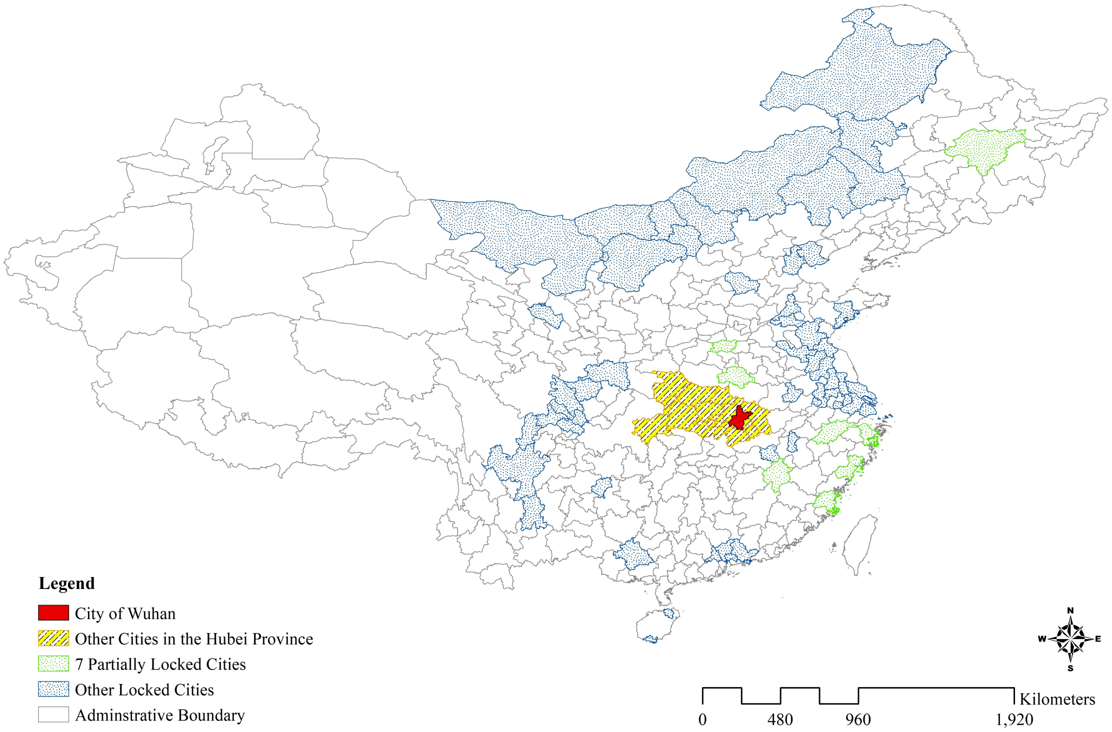
Notes: This table summarizes different levels of prevention and control measures across 115 cities. Panel A lists 17 cities with completed lockdown, which means all public transport and private vehicles are banned in the city, all residential buildings are locked down, and residents are not allowed to leave the city. 7 Cities in Panel B are under partial lockdown, the majority of the public transportation has been temporarily shut down, checkpoints have been set up to control the inflow population, and surveillance and tighter controls in each neighborhood. 91 Cities in Panel C set up checkpoints and quarantine zones, and public transport maintains normal operation.

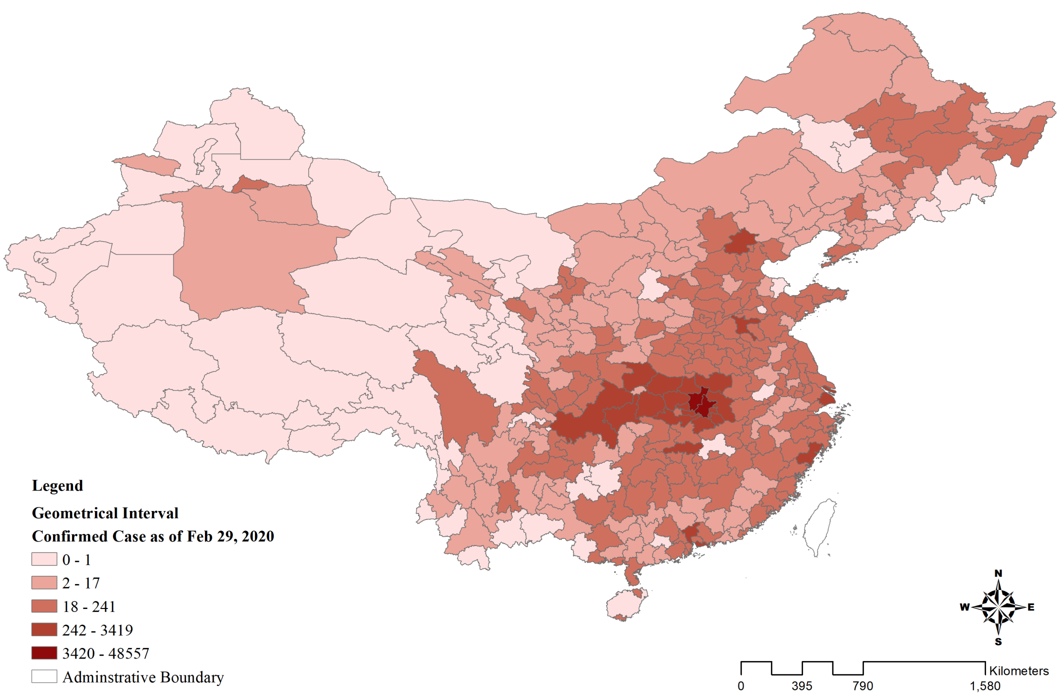
Table A3. Summary Statistics of Cities with Different Level of Controls

	Wuhan	7 Partial Lockdown Cities	Other Unlocked Cities
# of Total Cases as of Feb 2, 2020	5,142	101	13
Daily Average Population Inflow (2019 Sample Period)	466,682	294,799	78,006
Daily Average Population Outflow (2019 Sample Period)	420,900	282,995	78,789
Daily Average Within-City Population Flow (2019 Sample Period)	3,270,509	3,305,080	1,316,090
Permanent Population (2019)	9,785,388	8,433,975	3,524,548
GDP (Trillion CNY) (2019)	1,153	811	92

Notes: This table provides summary statistics on the count of total confirmed cases as of February 2, 2020, and on daily average population inflow, outflow, and within-city flow between January 12 and March 12 in 2019, and on permanent population GDP as of December 2019, for cities with different level of controls.

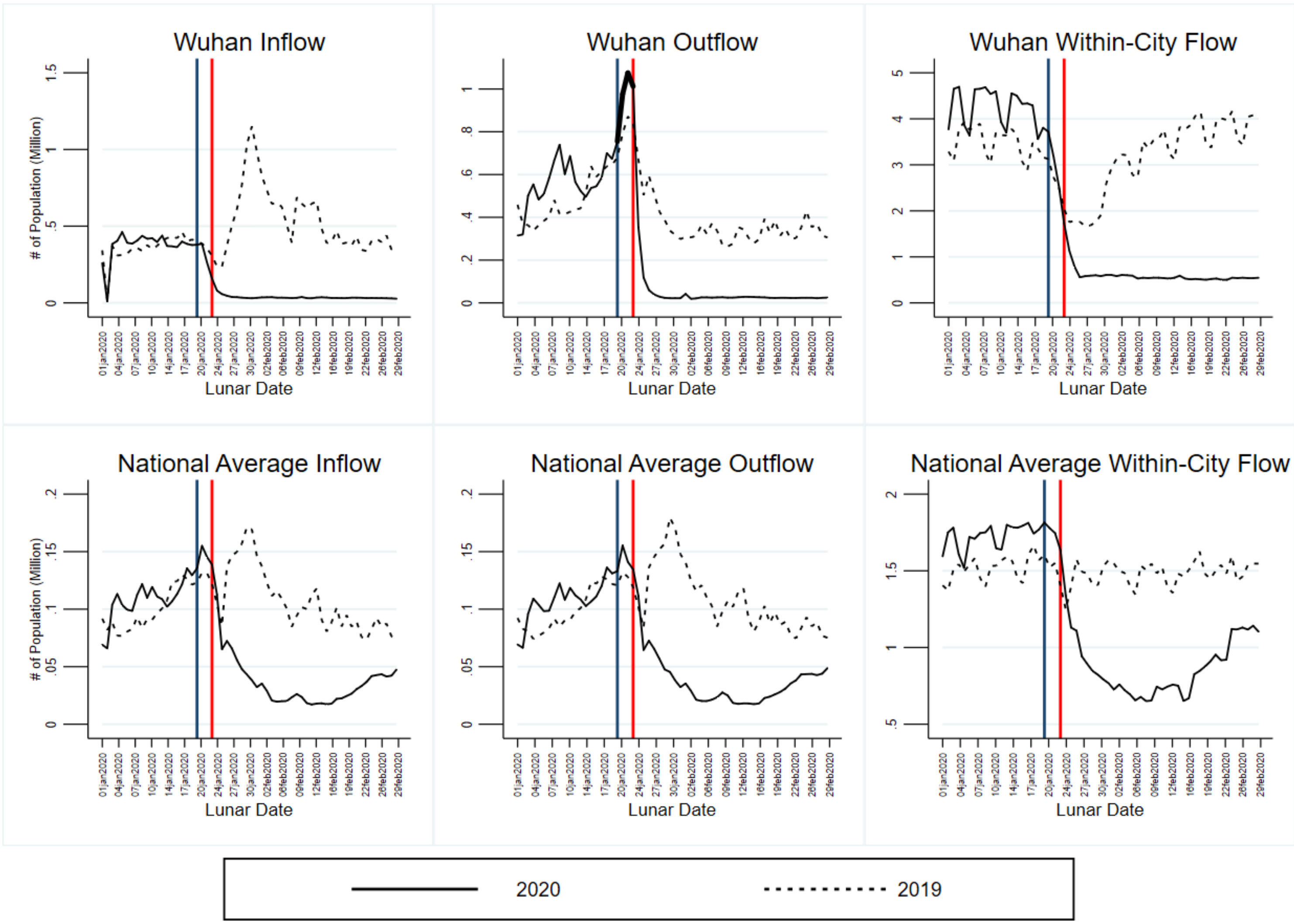
Figure A1. Cities with Control Measures and the Confirmed COVID-19 Cases





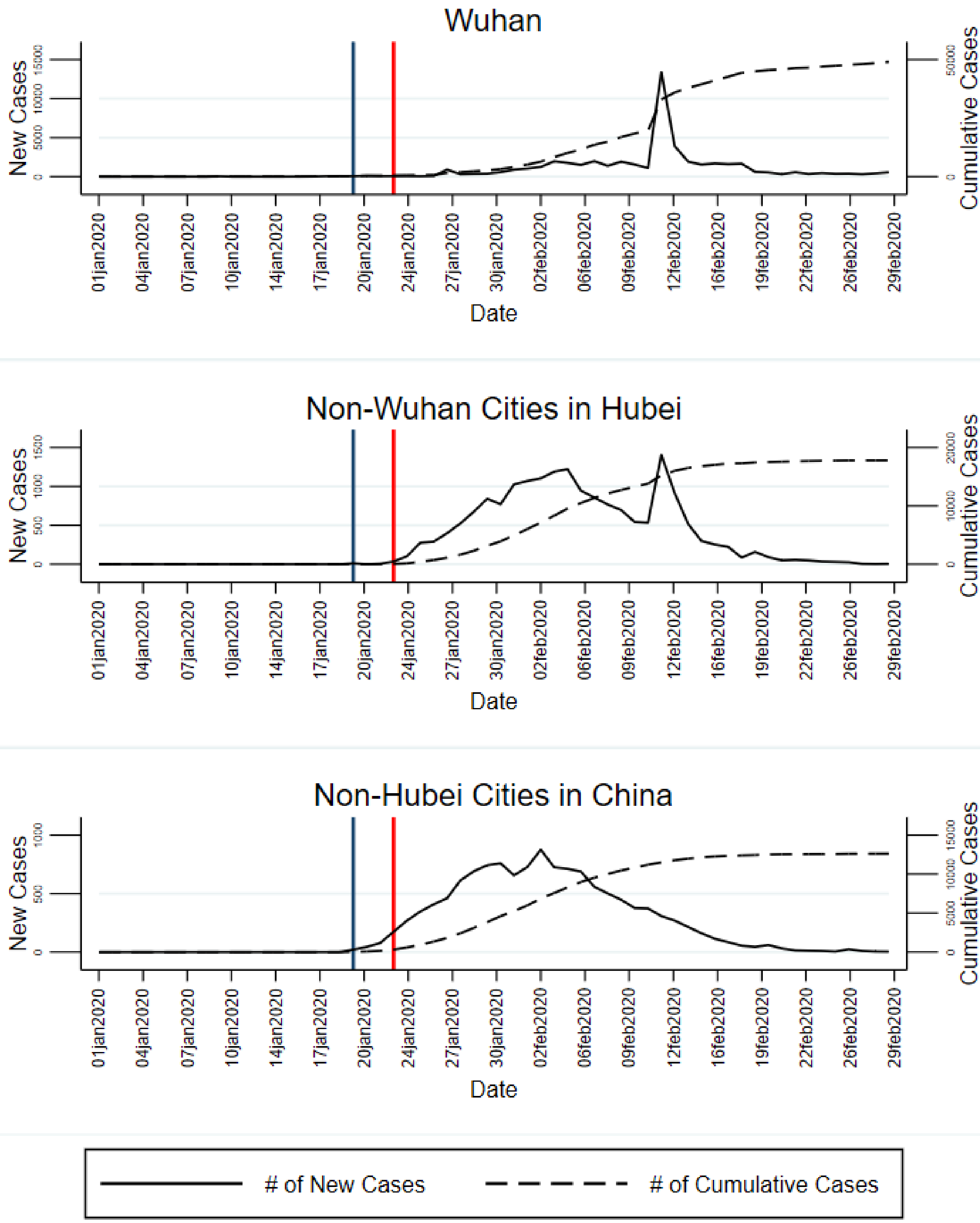
Notes: This figure presents the geographic distributions of cities with different levels of control measures and the number of Confirmed COVID-19 Cases as of Feb 29, 2020. The maps were plotted with ArcGIS 10.2 (ESRI).

Figure A2. Inter-city and Within-city Population Flows



Notes: Top figures show the inflows into Wuhan, outflows from Wuhan and within-Wuhan flows, for year 2020 (in the solid line) and year 2019 (in the dashed line), matched by the lunar calendar; and the bottom shows the corresponding figures for the national city averages. The blue vertical line indicates the date of January 20, 2020 when the health ministry confirmed human-to-human transmission of COVID-19; and the red vertical line indicates the date of January 23, 2020 when Wuhan was locked down.

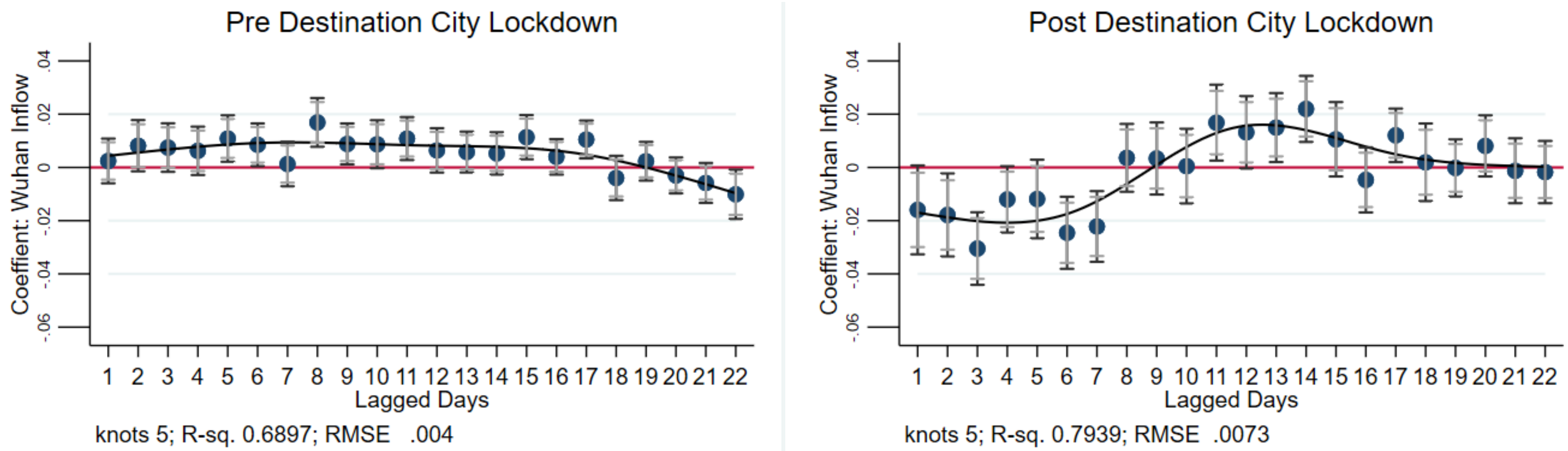
Figure A3. Daily Confirmed Cases and Cumulative Confirmed Cases



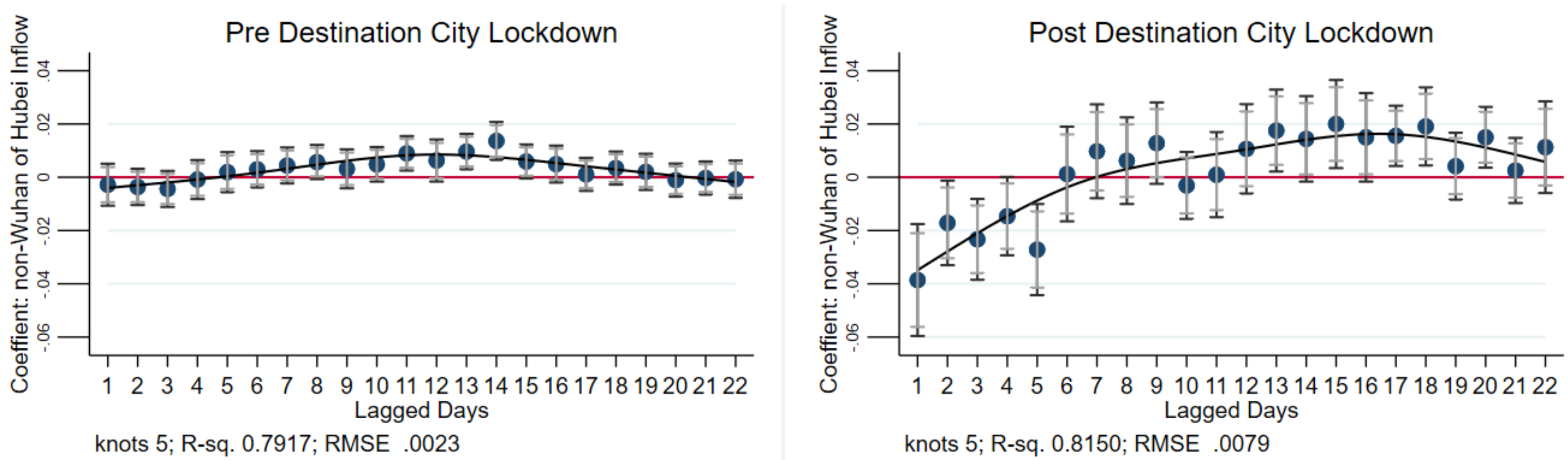
Notes: This figure shows the daily confirmed, dead, and healed cases in Wuhan, other cities in Hubei Province, and cities outside of Hubei Province.

Figure A4. Dynamic Impacts of Pre and Post Destination City Lockdown

Panel A: Inflows from Wuhan



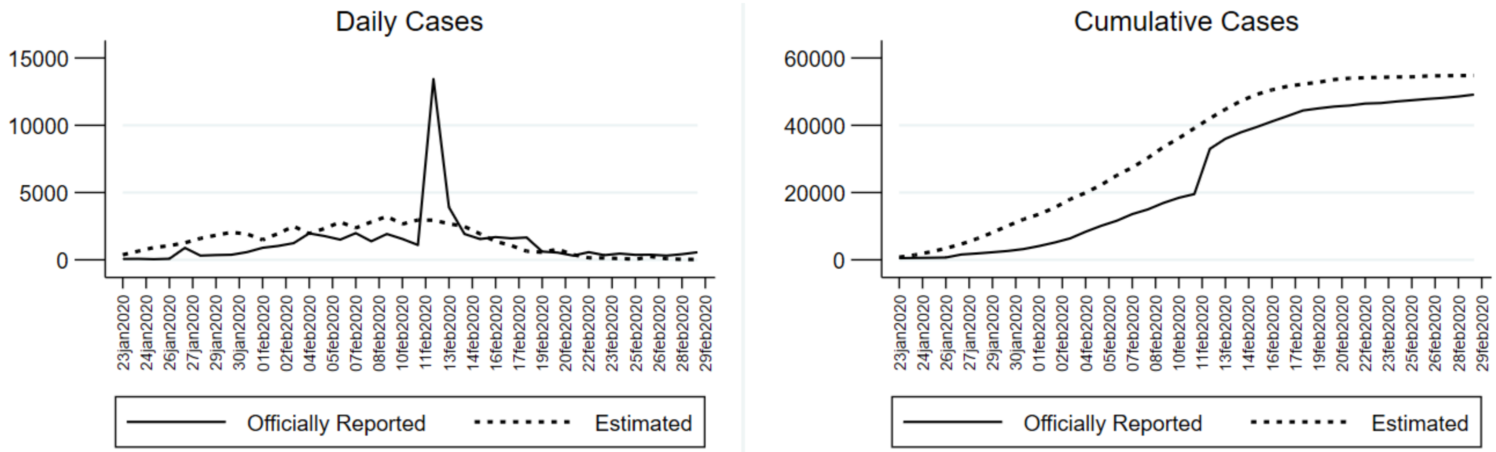
Panel B: Inflows from non-Wuhan Cities of Hubei



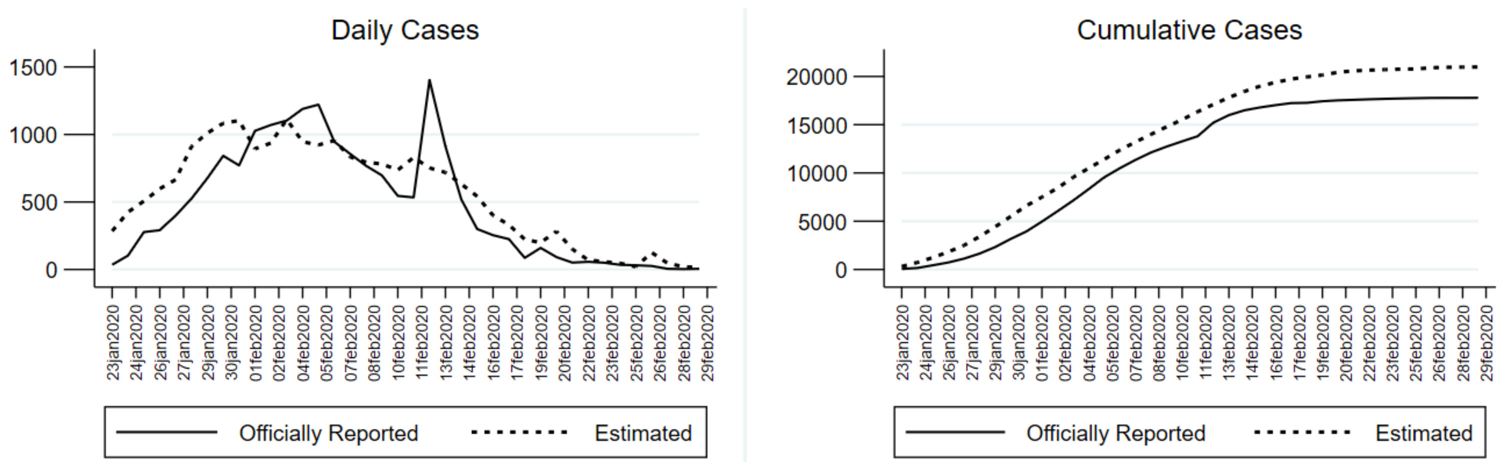
Notes: Panel A plots the dynamic lagged effects of past inflows from Wuhan pre (left figure) and post (right figure) destination cities' lockdown policy, if any. Panel B plots the dynamic lagged effects of past inflows from non-Wuhan Cities of Hubei pre (left figure) and post (right figure) destination cities' lockdown policy, if any. The coefficient estimates are obtained from estimating Equation (4). We add spline smoothing fit curves (in red) using the *rcspline* function and plot the 90% (the vertical gray whiskers) and 95% (the vertical black whiskers) confidence intervals.

Figure A5. Estimation on the “Actual” Number of Infected Cases in Wuhan and other Cities of Hubei

Panel A: Wuhan

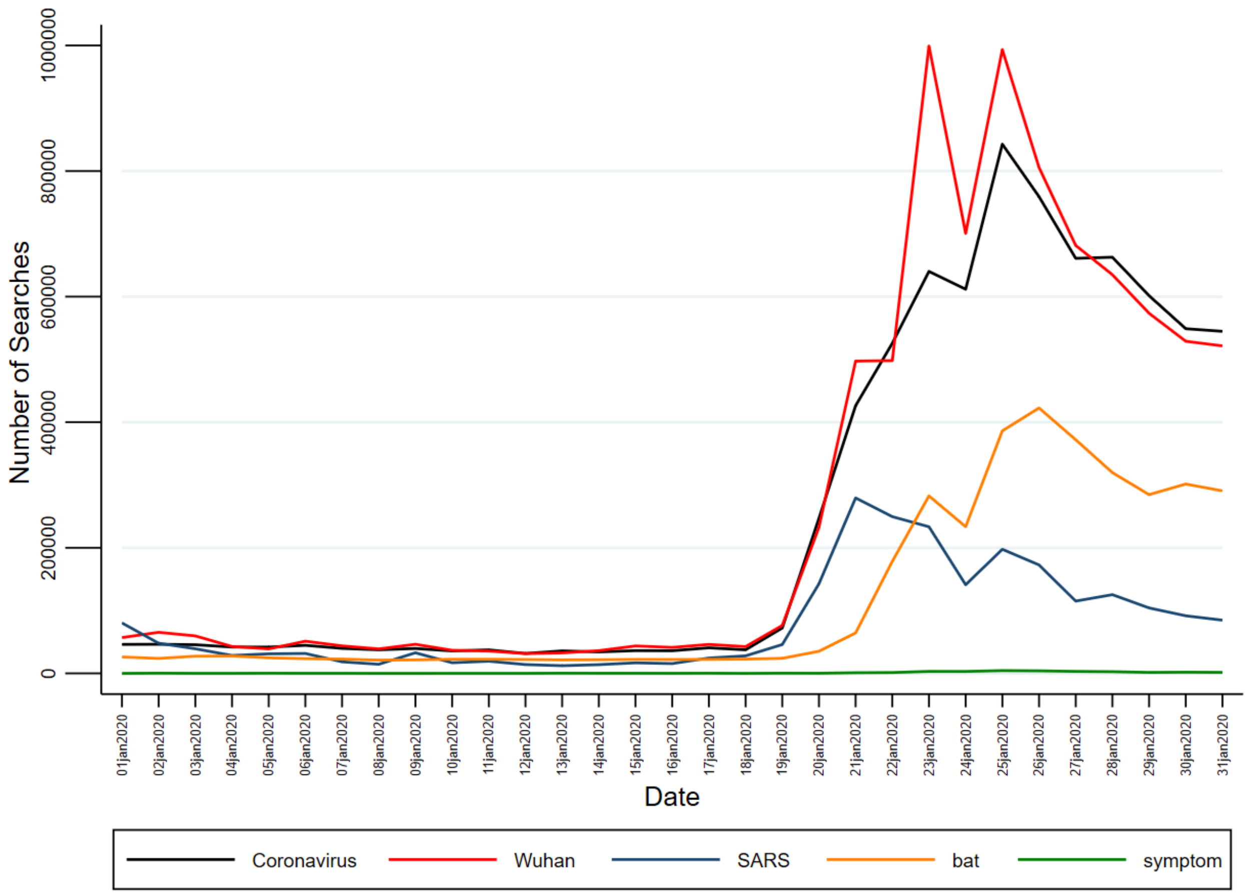


Panel B: 16 other Cities of Hubei



Notes: This figure compares the estimated COVID-19 cases (in the dotted curve) with the officially reported confirmed cases (in the solid curve) in Wuhan (top) and in 16 non-Wuhan cities in Hubei Province (bottom) from January 23 to February 29 in 2020. The left panel plots the estimated new COVID cases on each date t from 23 (January 23, 2020) to 60 (February 29, 2020). The right panel plots the estimated cumulative cases each day.

Figure A6: Daily Search Frequencies of COVID-19 Related Keywords



Notes: This figure graphs the daily search frequency of COVID-19 related keywords, including "Coronavirus", "Wuhan", "Bat", "SARS", and "symptom" (in Chinese). It shows a clear and abrupt spike in the search frequencies on January 20, the day of the public confirmation of human-to-human transmission of the Novel Coronavirus by Dr. Zhong Nanshan on Chinese national TV.

Appendix B. The Conversion of Baidu Mobility Indices into Number of Population Movements

We obtain three migration intensity indicators (the daily *in-migration* index (IMI) of a city, the daily *out-migration* index (OMI) of a city, and the daily *within-city* migration index (WCMI)) from Baidu Migration for 364 Chinese cities. Baidu Migration[16] uses Baidu Maps Location Based Service (LBS) open platform and Baidu Tianyan to calculate and analyze the LBS data, and provides visual presentation to show the trajectory and characteristics of the population migration. Baidu has been the dominant search engine in China because all Google search sites have been banned in mainland China since 2010.

Specifically, Baidu Migration provides the following information: (1), the top 100 origination cities (OC) for the population moving *into* the city and the corresponding percentages of inflow population that originated from each of the top 100 OC; and (2), the top 100 destination cities (DC) for the population moving *out of* the city, together with the corresponding percentages of the outflow population that go into each of the top DC. In the data, the cumulative percentages of the inflow population from the top 100 origination cities, and the cumulative percentages of the outflow population into the top 100 destination cities, reach 97% per city on average, which ensures that the Baidu Migration data capture near complete inflows and outflows for each of the 364 cities in the data.

We convert the mobility index unit into the number of population movements by taking the daily inflow of people into Shanghai by airplanes, trains, buses and cars, the daily within-city trips in Shanghai by subways, buses and expressways (provided by NESSDC), and the corresponding inflow index and within-city mobility index in Shanghai (provided by Baidu). Based on the definition of inter-city mobility indices provided by Baidu, the inter-city indices are comparable both across cities and time. We first divide the actual number of inflow/outflow population by the inflow/outflow index in a day to obtain population number per unit of inter-city index. For instance, given that the actual inflow population of 302,6000 on February 6, 2020 into Shanghai corresponds to the inflow index of 3.72, the population number per unit inflow index is 302,600/3.72=81344.08 on February 6, 2020. Since the NESSDC provides the actual number of inflow population between February 6 and February 22, 2020, we can then calculate population per unit inflow index for each day of this period and obtain an average population per inflow index, which equals to 90,848. To convert the inter-city flow indices into the total inter-city population flows for all cities, we multiply the indices by 90,848.

To convert the within-city mobility index to actual population flows, we weight the index by the number of the city's base population called"*regular residents*" in 2019 (i.e., people who had stayed in the city for at least six months during the year). For instance, given that the actual within-city population flows of 4,339,451 on February 6, 2020 in Shanghai corresponds to the within-city mobility index of 1.6 and the base population in Shanghai in 2019 is 24,237,800, the per unit within-city index is (4,339,451/24,237,800)/1.6)*24,237,800=2,712,156. Similarly, we calculate

[16]Source: `http://qianxi.baidu.com/`

the population per within-city mobility index for each day of this period to determine an average population per within-city flow index, which is 2,182,264. We then convert the within-city mobility indices into the number of within-city population flows for all cities by multiplying the indices by 2,182,264 and the ratio of a city's base population in 2019 over Shanghai's base population in 2019, 24,237,800. For instance, if the within-city mobility index in Wuhan on February 6, 2020 is 0.6 and Wuhan's base population in 2019 is 9,785,388, then the actual number of within-city population flows in Wuhan on February 6, 2020 is 2,182,264*0.6*(9,785,388/24,237,800)=528,620.

Appendix C. Estimating the "Actual" Number of Infection Cases in Wuhan and Other Cities in Hubei

Anecdotal evidence suggests the official statistics of COVID-19 cases in Wuhan may have been under-reported due to the shortages of testing equipment and other medical resources. With the estimated dynamic effects shown in Figure 1, which is estimated under the plausible assumption that the reported cases outside Hubei are reliable, we can estimate the "actual"' number of infection cases in Wuhan and other cities in Hubei.

To estimate the "actual" number of infections in Wuhan using the estimates from Eq. (3), we technically need to impute a value for $\mathit{Inflow}_{WH,WH,t-\kappa}$, that is "inflows from Wuhan to Wuhan." We proxy these inflows by the daily *within-Wuhan* population movement from January 1 to February 29, i.e., by $\mathit{WithinCityFlow}_{WH,t-\kappa}$. Similarly, to estimate the "actual" number of infection cases in other cities in Hubei, we need to replace the inflow from Wuhan to itself by the corresponding daily within-city-j population movements.

However, we do not have the city fixed effects for cities in Hubei because they were not included in the estimation sample for Eq. (3). Therefore, we cannot directly predict the "actual" number of infections in Wuhan and other cities in Hubei using Eq. (6). Instead, we use the estimated $\hat{\beta}_{1\kappa}$ coefficient as a measure of the *elasticity* of the new cases outside of Hubei at date t with respect to inflows from Wuhan κ days ago. We then calculate the percentage difference between the within-Wuhan population flow $\mathit{Inflow}_{WH,WH,t-\kappa}$ and the average inflows from Wuhan to cities outside Hubei, i.e., $\left(\sum_{i=1}^{347} \mathit{Inflow}_{i,WH,t-\kappa}\right)/347$, together with the average new daily cases outside of Hubei at date t and $\hat{\beta}_{1\kappa}$, to impute what Wuhan new cases would have been at date t, *under the assumption that* the relationship dictating the within-Wuhan population movements and Wuhan's "actual" new cases at date t is similar to that estimated for cities outside Hubei in Eq. (3). We follow the same method to estimate the "actual" number of new infections in 16 other cities of Hubei.

In Appendix Figure A5, we plot the estimated daily new cases according to the above-described method using the estimated Eq. (3), as well as the corresponding cumulative cases for Wuhan (Panel A) and 16 other cities of Hubei (Panel B) for the period of January 23 to February 29, 2020. We also plotted the corresponding daily and cumulative officially reported (i.e., documented) cases.

We find a persistent gap between the estimated and reported laboratory-confirmed cases in Wuhan before February 11, 2020, just before the announcement of a new Party Secretary for Hubei on February 12, 2020. The estimated "actual" number of infection cases is 2.81 times the reported cases during the first 20 days after the Wuhan lockdown, on average. In particular, we estimate that on January 23, 2020, the day of the Wuhan lockdown, 38.29% of our estimated infections in Wuhan were undocumented in the sense that the number of officially reported cases on that day was only 61.71% of our estimated infection cases. This undocumented-real gap widened over time, possibly due to the overwhelmed health care system, and peaked at 79.57% on January 26. The proportion of undocumented infections started to decline gradually, when more medical support and resources were mobilized across China to support Wuhan. As of February 29, we estimate that there were 54,797 total COVID-19 infections in Wuhan, which is 11.55% higher than the official reported statistics for Wuhan - a total of 49,122 cases. The 11.55% discrepancy can be plausibly be explained by the unaccounted for self-healing and death that might have occurred during the early periods of the outbreak between January 23 and early February. Thus, we are led to conclude that the almost all infection cases in Wuhan were able to be treated over time as the stress on the health system was relieved, and moreover, the official statistics were mostly accurate, as can be seen from the left figure on the daily new cases in Panel A in Figure A5.

In the bottom panel of Figure A5, we plot our estimated daily new confirmed cases and total infection case for 16 cities (other than Wuhan) in Hubei, together with the officially reported series. We find that in the 16 cities, infections were more seriously under-reported in the first week after the Wuhan lockdown when our estimated infected cases are 1.87 times the reported cases. Our estimate reveals a very high rate of undocumented infections on the first day of Wuhan lockdown: 81.02%. The gap narrowed gradually with more medical resources provided and more stringent control measures implemented in those cities. By the end of our study period on February 29, 2020, the estimated "actual" number of infections is 20,981 cases in 16 other cities in Hubei, which is 17.97% higher than the officially reported cumulative cases (17,785). The discrepancy between the estimated and officially reported cumulative cases could at least be partially attributed to the unaccounted for self-healing and death that might have occurred during the early periods of the outbreak.