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

Hanming Fang[1,2], Long Wang[2], Yang (Zoe) Yang[3]

[1]University of Pennsylvania & NBER

[2]ShanghaiTech University

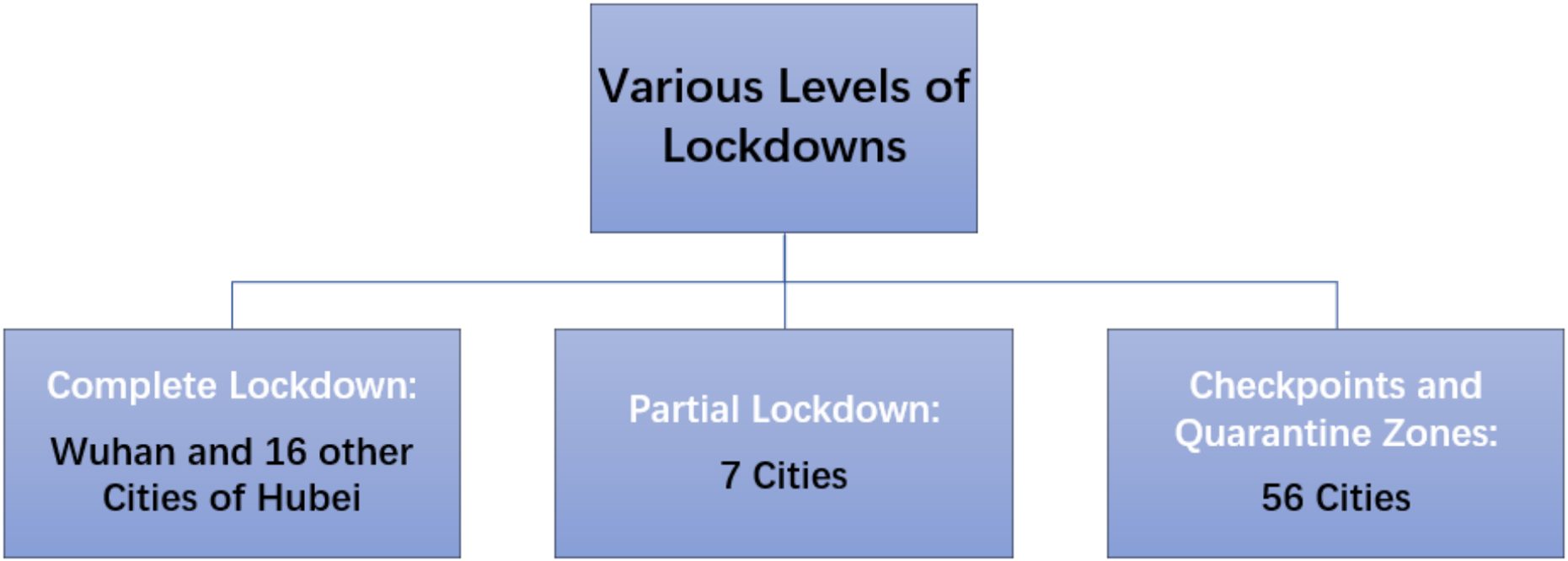
[3]The Chinese University of Hong Kong

May 1, 2020

Key Known Information of 2019-nCoV

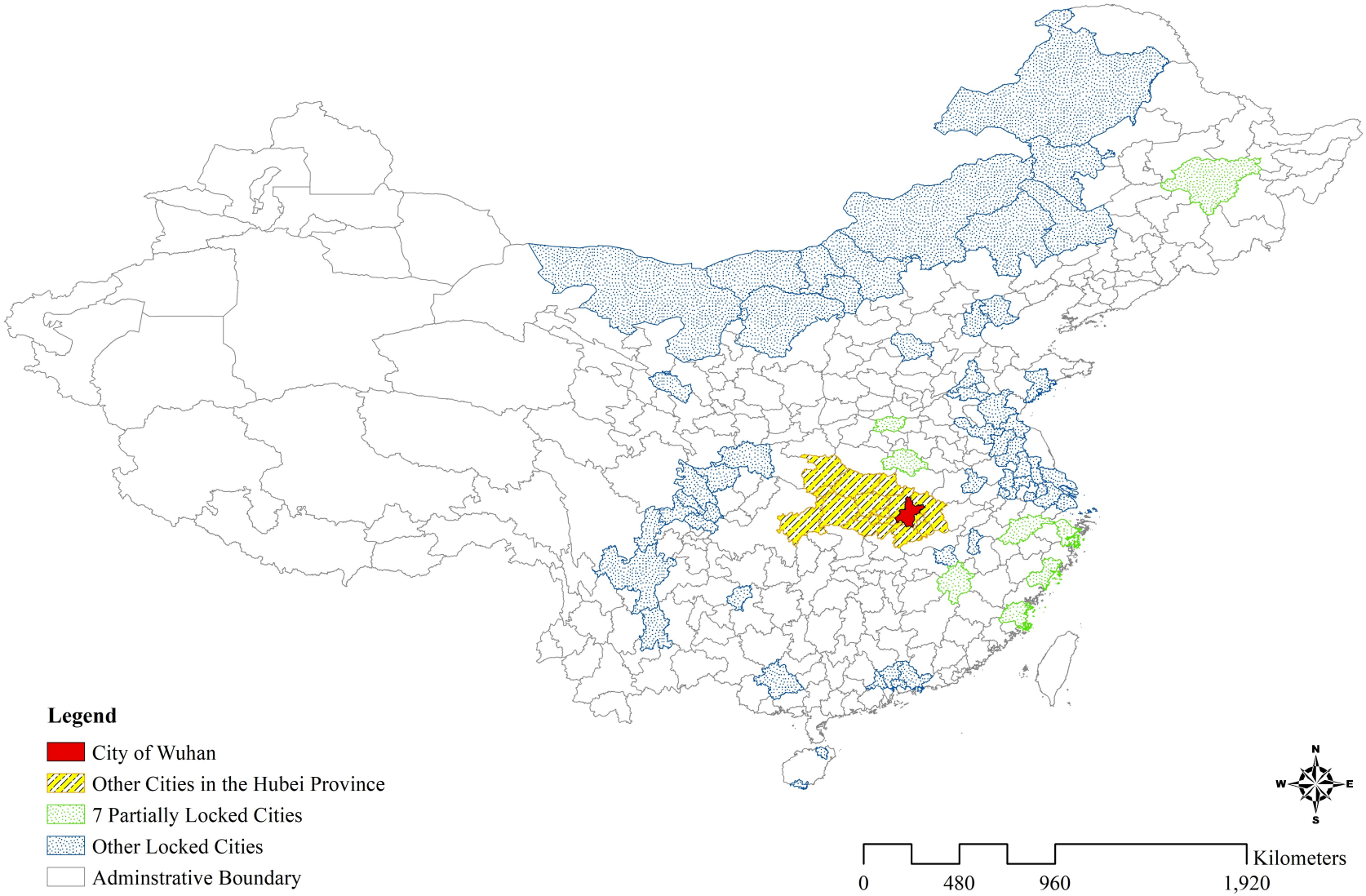
- First detected in Wuhan, China, with the first case reported in early to mid December 2019 ("pneumonia with unknown origin")
- Common symptoms at onset of illness were fever, cough, and myalgia or fatigue (Huang et al., 2020, *Lancet*)
- China publicly confirmed human-to-human transmission on **Jan 20, 2020**
- A mean of 3.28 and a median of 2.79 of the basic reproduction number (R0) (Li et al., 2020, *Journal of Travel Medicine*)
- Transmission from an asymptomatic contact (Rothe et al., 2020, *New England Journal of Medicine*)
- Median incubation period was 4 days (interquartile range, 2 to 7) (Guan et al., 2020, *New England Journal of Medicine*)
- Ranges from 2 to 14 days, or even as long as 24 days (Lauer et al., 2020, Annals of Internal Medicine)
- Currently, no licensed vaccines or specific therapeutics (*US Food & Drug Administration, 2020*)

Background: Lockdowns of Cities in China



- An unprecedented *cordon sanitaire* - strict lockdown, of the epicenter from **10am on Jan. 23, 2020**

Background: Geographic Distribution of Lockdown Cities in China



Research Questions

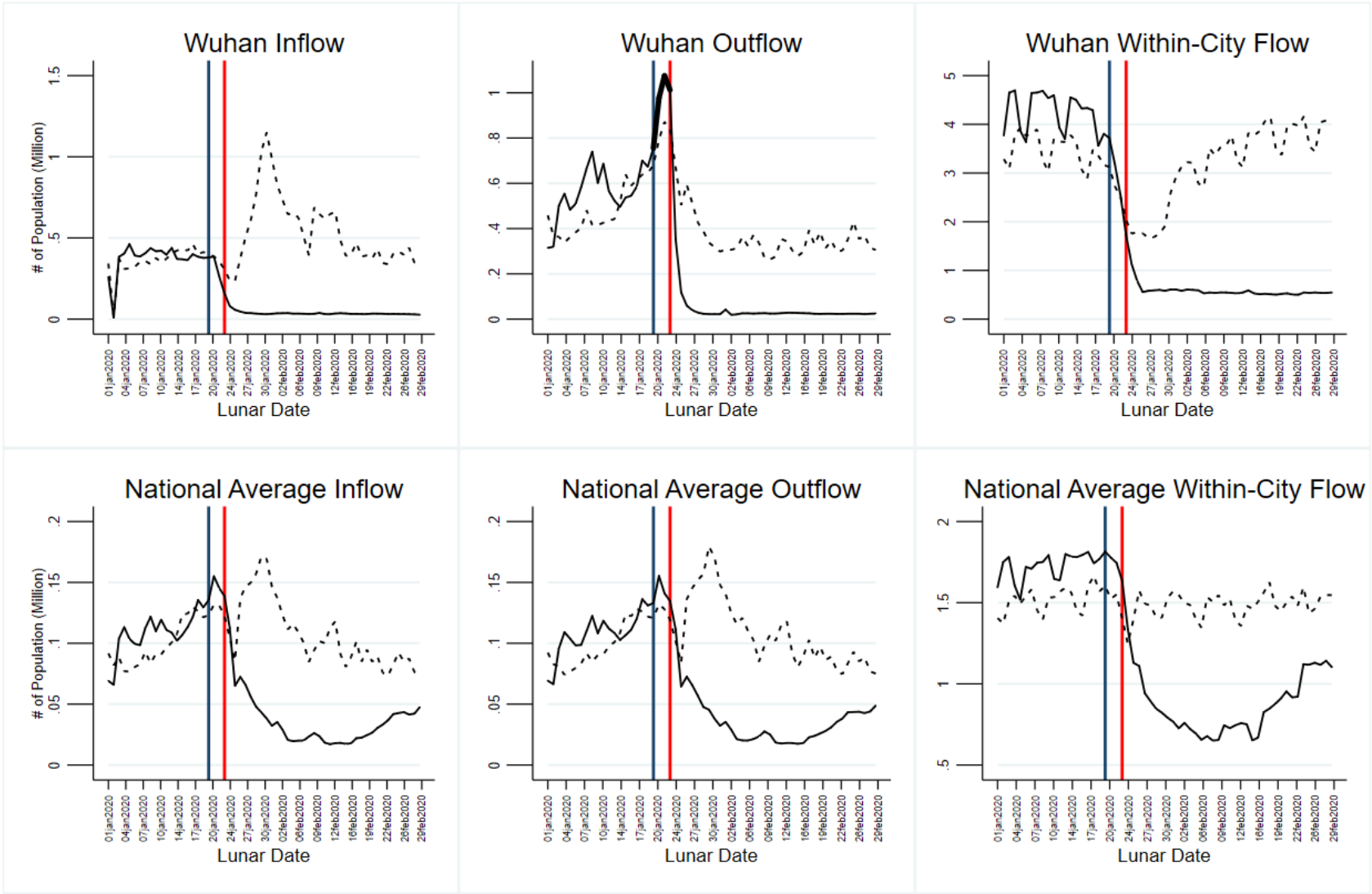
1. How does the Wuhan lockdown affect population movement?
2. How do outflows from Wuhan and other cities in Hubei province affect virus infection in the destination cities?
3. What are the "actual" numbers of COVID-19 cases in Wuhan and other cities in Hubei?
4. How many COVID-19 cases elsewhere in China were prevented by the unprecedented Wuhan lockdown?
5. Are social distancing policies in destination cities effective in reducing the spread of the infections?

Data

- Population Migration Data from Baidu
  - data period: between Jan 12 and Mar 12 in 2019, and between Jan 1 and Feb 29 in 2020
  - in-migration index, out-migration index, and within-city-migration index
  - Convert the index to the number of people using date from National Earth System Science Data Center collected and reported by Shanghai from February 1, 2020
    - ★ 90,848 person movements per inter-city index unit
    - ★ 2,182,264 person movements per within-city index unit for the city of Shanghai (for other cities, we need to scale it to their base populations relative to Shanghai's population)
  - 5,955,798 city-pair-day observations for 120,142 pairs of cities for 364 Chinese cities
  - 43,310 city-day level observations for within-city mobility

- COVID-19 Data from China CDC
  - data period: between Jan 11 and Feb 29, 2020
  - daily updates on confirmed, dead, and recovered COVID-19 cases in 296 cities

Background: Inter-City and Within-City Population Mobility



Question 1: What is the causal impact of Wuhan lockdown on population movements?

- **Challenges** in identifying the *pure lockdown* effects
  - confounds with the Spring Festival effect
  - confounds with the virus effect
    - ★ in the absence of lockdown, people attempt to avoid exposure to the virus in the journeys and public spaces
    - ★ applies everywhere
  - confounds with the panic effect
    - ★ in the absence of lockdown, people attempt to flee from the epicenter, and avoid entering the epicenter
    - ★ specific to the epicenter
- **Strategies** in disentangling these effects
  - create a specific pre-lockdown period between Jan 20 and Jan 22, 2020 to capture the panic effect
    - ★ confirmation of human-to-human transmission on Jan 20
  - employ several difference-in-differences (DID) estimation specifications by comparing different treatment and control groups to estimate virus and lockdown effects

The impact of Wuhan lockdown on inter-city population movements

- The DID specification for **inter-city** population mobility:

$$\begin{aligned} Ln(Flow_{i,j,t}) =& \alpha + \beta_1 \cdot Treat * Before_{1,t} + \beta_2 \cdot Treat * Before_{2,t} \\ &+ \beta_3 \cdot Treat * After_t + \mu_{i,j} + \theta_t + \epsilon_{i,j,t} \end{aligned} \tag{1}$$

  - $Ln(Flow_{i,j,t})$, is the logarithmic population flows received by city i from city j at date t
  - $Before_{1,t} = 1$ for the period from Jan 11 to Jan 19, 2020
    - ★ used to test the parallel trend assumption
  - $Before_{2,t} = 1$ for the period from Jan 20 to Jan 22, 2020
    - ★ used to examine the panic effect
  - $After_t = 1$ for the period between Jan 23 and Feb 29, 2020
  - The city-pair fixed effect $\mu_{i,j}$ and the date-fixed effect θ_t are included
  - The standard errors are clustered at the date level

The impact of Wuhan lockdown on within-city population movements

- The DID specification for **within-city** population mobility:

$$\begin{aligned} Ln(WithinCityFlow_{i,t}) =& \alpha + \beta_1 \cdot Treat * Before_{1,t} + \beta_2 \cdot Treat * Before_{2,t} \\ &+ \beta_3 \cdot Treat * After_t + \mu_i + \theta_t + \epsilon_{i,t} \end{aligned} \tag{2}$$

  - $Ln(WithinCityFlow_{i,t})$ is the logarithmic within-city population mobility measure for city i at date t
  - $Before_{1,t}$, $Before_{2,t}$ and $After_t$ are defined in the same way as in Equation (1)
  - City fixed effects μ_i and date fixed effects θ_t are included
  - The standard errors are clustered at the date level

- The definition of **Treat** varies by specific DID designs, and we will be explicit about its definition in the result section

Results: Impact of Lockdown on Outflow

Treatment Group	Wuhan 2020	Wuhan 2020	Wuhan 2020
Control Group	284 Unlocked Cities 2020	Wuhan 2019	7 Cities 2020
Effects	Lockdown + Virus + Panic	Lockdown + Virus +Panic	Lockdown
Model	(1)	(2)	(3)
Panel B. Dep. Variable: *ln(Outflow Population)*			
Treat*Before$_1$	-0.09	0.045	0.01
	(0.059)	(0.079)	(0.018)
Treat*Before$_2$	0.008	0.728***	0.174***
	(0.071)	(0.071)	(0.03)
Treat*After	-1.364***	-1.287***	-0.829***
	(0.141)	(0.152)	(0.035)
Observations	2,184,401	26,542	71,533
R-squared	0.825	0.898	0.901

- **Model 1:** Wuhan 2020 (Treat) vs. 284 Unlocked Cities 2020 (Control)
- **Model 2:** Wuhan 2020 (Treat) vs. Wuhan 2019 (Control)
- **Model 3:** Wuhan 2020 (Treat) vs. Seven Other Lockdown Cities 2020 (Control)

Results: Impact of Lockdown on Inflow

Treatment Group	Wuhan 2020	Wuhan 2020	Wuhan 2020
Control Group	284 Unlocked Cities 2020	Wuhan 2019	7 Cities 2020
Effects	Lockdown + Virus + Panic	Lockdown + Virus +Panic	Lockdown
Model	(1)	(2)	(3)
Panel A. Dep. Variable: *ln(Inflow Population)*			
Treat*Before$_1$	-0.088	-0.149	-0.023
	(0.143)	(0.152)	(0.021)
Treat*Before$_2$	-0.005	-0.122	-0.019
	(0.143)	(0.083)	(0.039)
Treat*After	-2.125***	-2.522***	-1.454***
	(0.245)	(0.208)	(0.043)
Observations	2,174,711	26,824	72,488
R-squared	0.825	0.805	0.763

Results: Impact of Lockdown on Within-city Flow

Treatment Group	Wuhan 2020	Wuhan 2020	Wuhan 2020
Control Group	284 Unlocked Cities 2020	Wuhan 2019	7 Cities 2020
Effects	Lockdown + Virus + Panic	Lockdown + Virus +Panic	Lockdown
Model	(1)	(2)	(3)
Panel C. Dep. Variable: *ln(Within-city Population Flow)*			
Treat*Before$_1$	-0.004	-0.019	0.044
	(0.088)	(0.063)	(0.042)
Treat*Before$_2$	-0.323**	-0.277***	-0.094
	(0.123)	(0.064)	(0.101)
Treat*After	-1.339***	-1.871***	-0.780***
	(0.09)	(0.089)	(0.129)
Observations	17,198	122	256
R-squared	0.976	0.952	0.936

Results: Summarizing the Panic Effect, Virus Effect and Lockdown Effect

Effect	Infows	Outflows	Within-City
Panic Effect	-11.48%	+107.09%***	-24.19%***
Virus Effect	-65.63%***	-36.75%***	-66.41%***
Lockdown Effect	-76.64%***	-56.35%***	-54.15%***

Question 2: What is the impact of Lockdown on the National Spread of COVID-19?

- **Motivation**: inflows from Wuhan with different lags may have differential impacts on the current new cases in the destination cities
- A dynamic distributed lag regression:

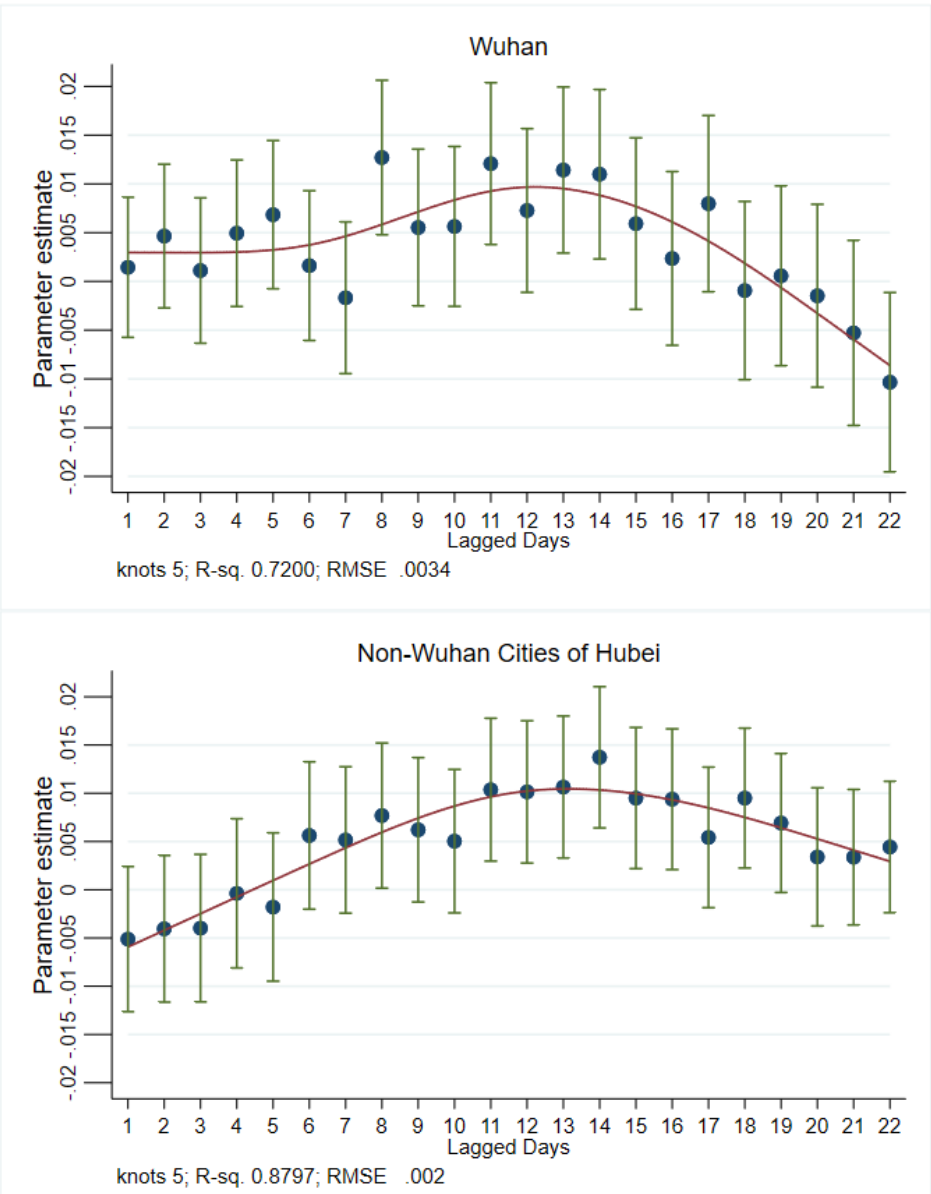
$$Ln(1+NewCase_{i,t}) = \alpha + \sum_{\kappa=1}^{22} \beta_{1\kappa} \cdot Ln\left(Inflow_{i,WH,t-\kappa}\right) + \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa}\right) + \mu_i + \theta_t + \epsilon_{it} \quad (3)$$

  - i indexes the cities outside of Hubei, and $t \in \{23, ..., 60\}$ indicates the date
  - $\kappa \in \{1, ..., 22\}$ indicates the time lapsed from the inflows from Wuhan or other Hubei cities till the current date t
  - $Ln(1+NewCase_{i,t})$ is the logarithm of the number of new confirmed cases in city i at date t
  - $Inflow_{i,WH,t-\kappa}$ and $\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa}$ are the inflows from Wuhan, and the inflows from the 16 other cities in Hubei to city i, respectively
  - City fixed effects μ_i and date fixed effects θ_t are included
  - The standard errors are clustered at the date level

Data Concerns for Confirmed Cases in Hubei

- Be cautious about the confirmed cases in Hubei
  - lack of medical resources in the early phases of the virus outbreak
  - lack of incentives in reporting the "actual" number of confirmed cases
- asymptomatic cases may not be tested and confirmed: this is an issue for both cities inside and outside of Hubei

Results: Impact of Lagged Inflow on Current Cases

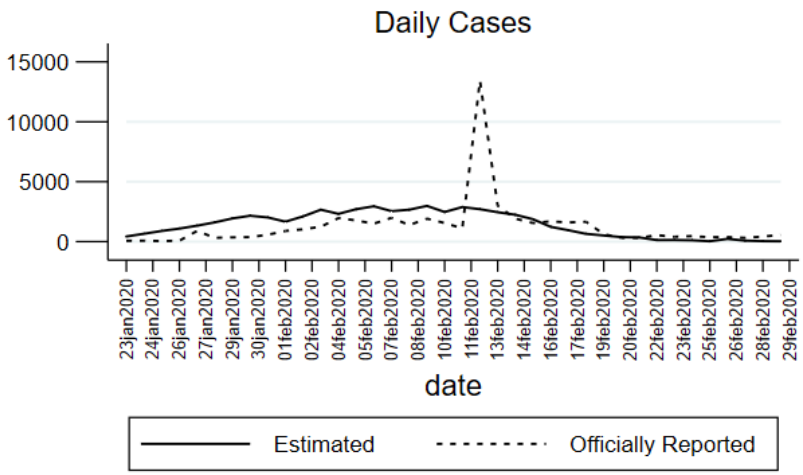


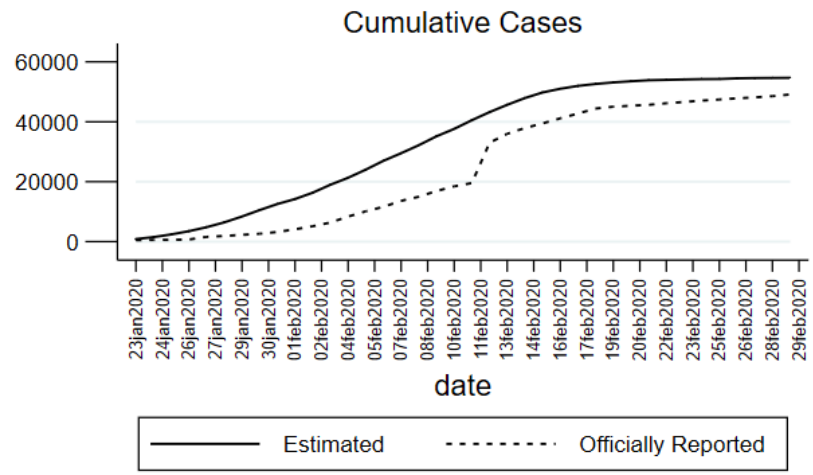
Question 3: What are the "Actual" Numbers of Cases in Hubei Cities?

- Assume the reported cases outside Hubei Province are reliable
- Use the estimated dynamic effects as shown in the left Figure to estimate the "actual" number of infections
- Use the within-Wuhan population movement as a proxy for "inflows from Wuhan to Wuhan"
- Use the within-Hubei-city-i population movement as a proxy for "inflows from Hubei city i to i"

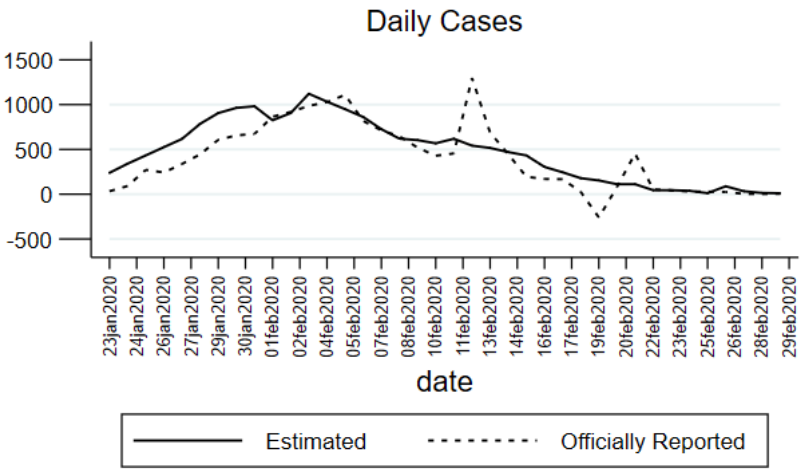
Estimating the "Actual" Number of Cases in Hubei Cities

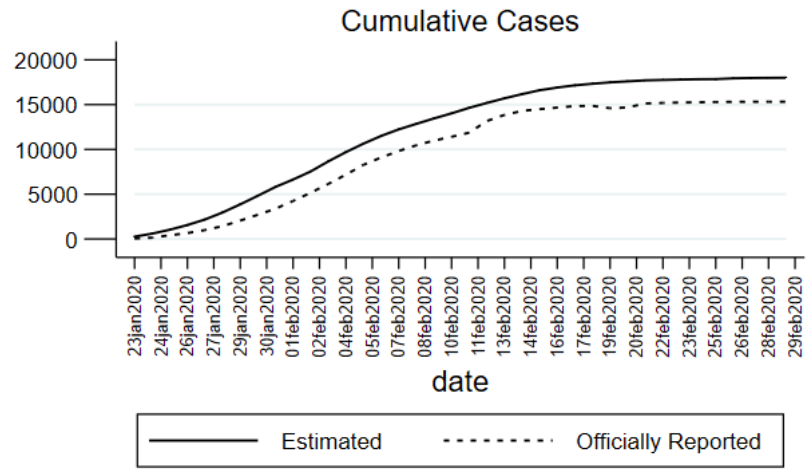
Panel A: Wuhan





Panel B: Non-Wuhan Cities of Hubei





Question 4: How many COVID-19 cases were actually prevented by the Wuhan lockdown in China?

- We simulate the counterfactual number of COVID-19 cases:

$$Ln(1+\widetilde{NewCase}_{i,t}) = \hat{\alpha} + \sum_{\kappa=1}^{22} \hat{\beta}_{1\kappa} \cdot Ln\left(\widetilde{Inflow}_{i,WH,t-\kappa}\right) + \sum_{\kappa=1}^{22} \hat{\beta}_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa}\right) + \hat{\mu}_i \quad (4)$$

  - $\hat{\beta}_{1\kappa}$ and $\hat{\beta}_{2\kappa}$ are coefficient estimates obtained from regressions specified in Equation (3)
  - we can predict the counterfactual COVID-19 cases without Wuhan lockdown if we know the *counterfactual* inflows from Wuhan to city i for days after Jan 23
  - City fixed effects μ_i and date fixed effects θ_t are included
  - The standard errors are clustered at the date level

Results: Estimating the Counterfactual Number of Cases

Effect	Infows	Outflows	Within-City
Panic Effect	-11.48%	+107.09%***	-24.19%***
Virus Effect	-65.63%***	-36.75%***	-66.41%***
Lockdown Effect	-76.64%***	-56.35%***	-54.15%***

- In the absence of Wuhan lockdown, we would expect that outflows from Wuhan in days after Jan 23 to be:

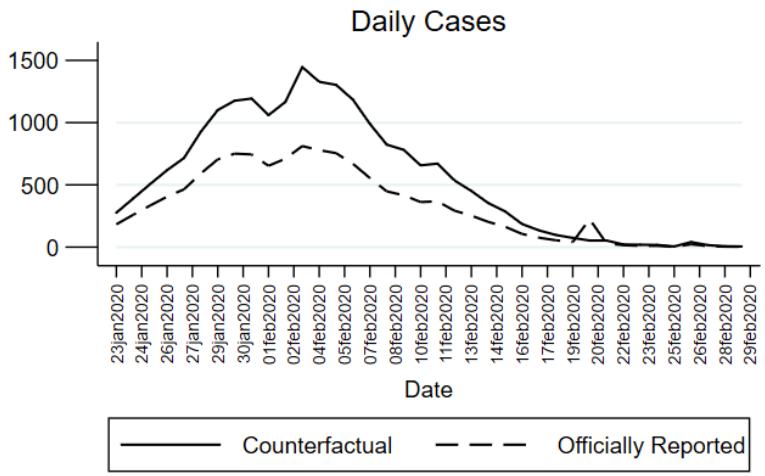
$$\overbrace{(1-0.3675)}^{\text{Reducton from Virus Effect}} * \overbrace{(1+1.0709)}^{\text{Increase from Panic Effect}} = 1.31$$

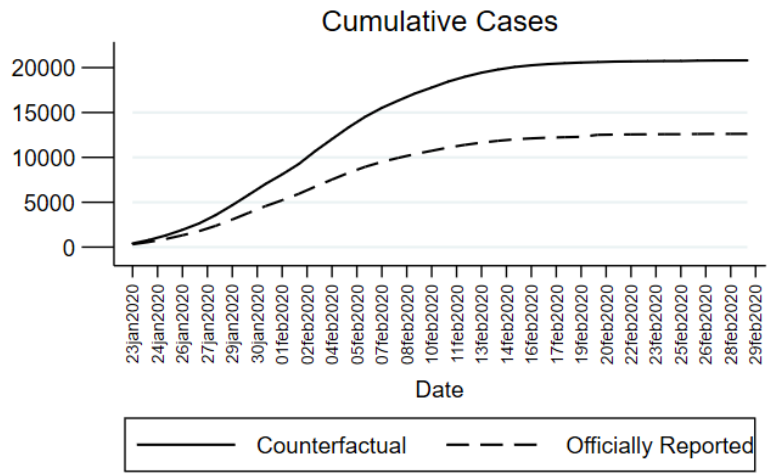
  times higher than the normal outflows from Wuhan to other cities.

- Using the counterfactual Wuhan inflow and the coefficient estimates obtained from Equation (3), we estimate the counterfactual number of cases to be:

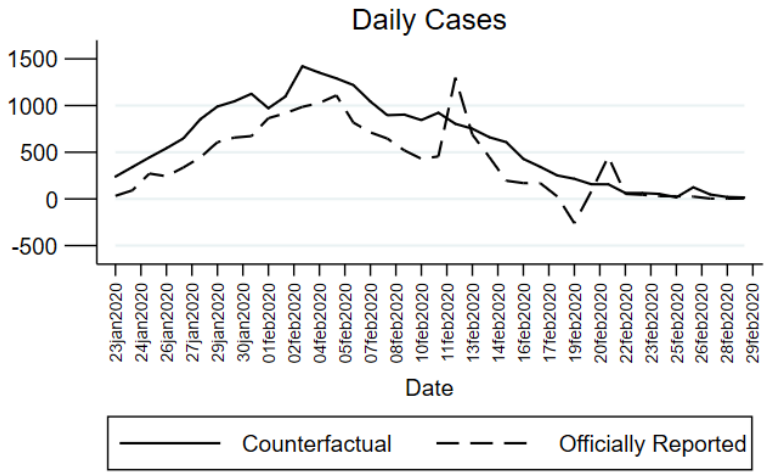
Results: Estimating the Counterfactual Number of Cases

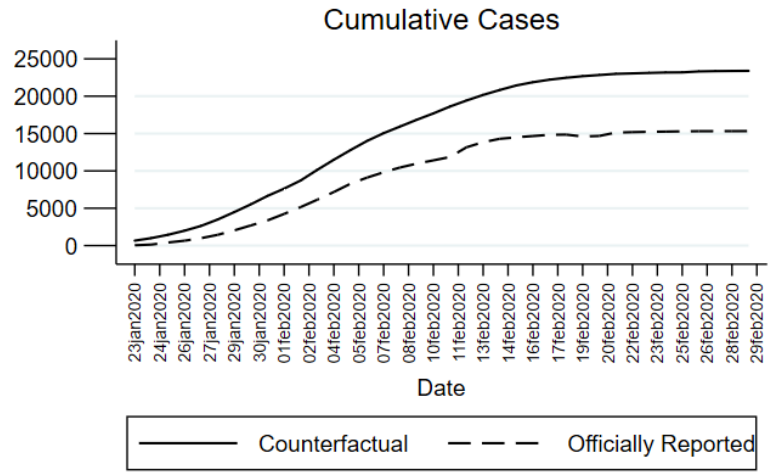
Panel A: 347 Cities outside Hubei





Panel B: 16 Non-Wuhan Cities Inside Hubei





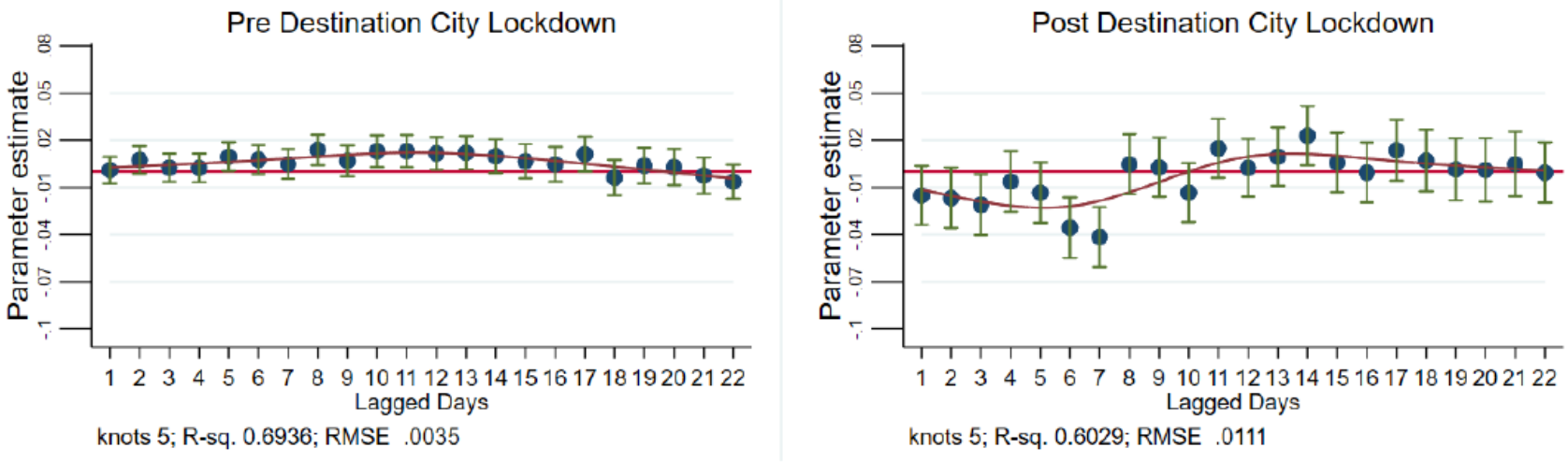
Question 5: What is the effect of social distancing on virus transmission?

- Specification:

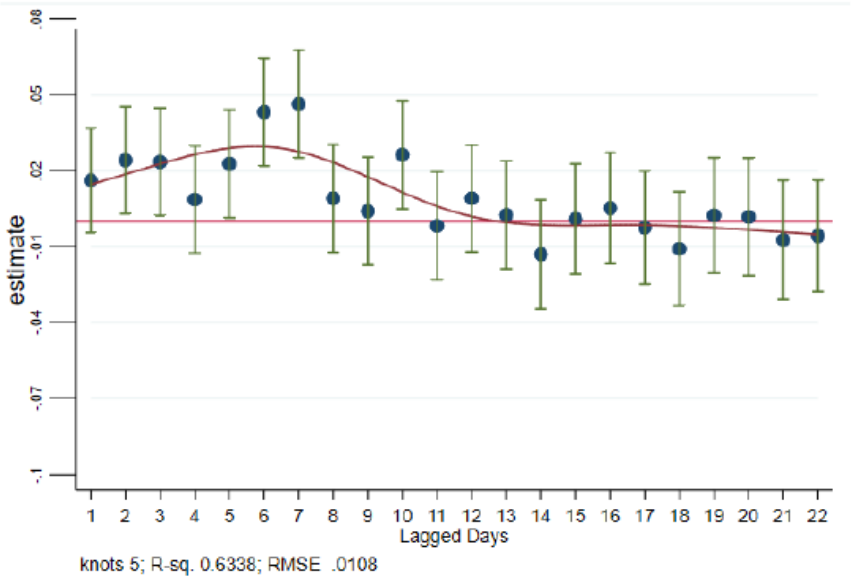
$$
\begin{aligned}
Ln(1+NewCase_{i,t}) = \alpha &+ \sum_{\kappa=1}^{22} \beta_{1\kappa} \cdot Ln\left(Inflow_{i,WH,t-\kappa}\right) \cdot \left(1 - Lockdown_{i,t}\right) \\
&+ \sum_{\kappa=1}^{22} \gamma_{1\kappa} \cdot Ln\left(Inflow_{i,WH,t-\kappa}\right) \cdot \left(Lockdown_{i,t}\right) \\
&+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa}\right) \cdot \left(1 - Lockdown_{i,t}\right) \\
&+ \sum_{\kappa=1}^{22} \gamma_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa}\right) \cdot \left(Lockdown_{i,t}\right)
\end{aligned}
\tag{5}
$$

  - $Lockdown_{i,t} = 1$ if time t is a date after destination city i's "lockdown" date

Results: Effect of Social Distancing on Wuhan Inflow

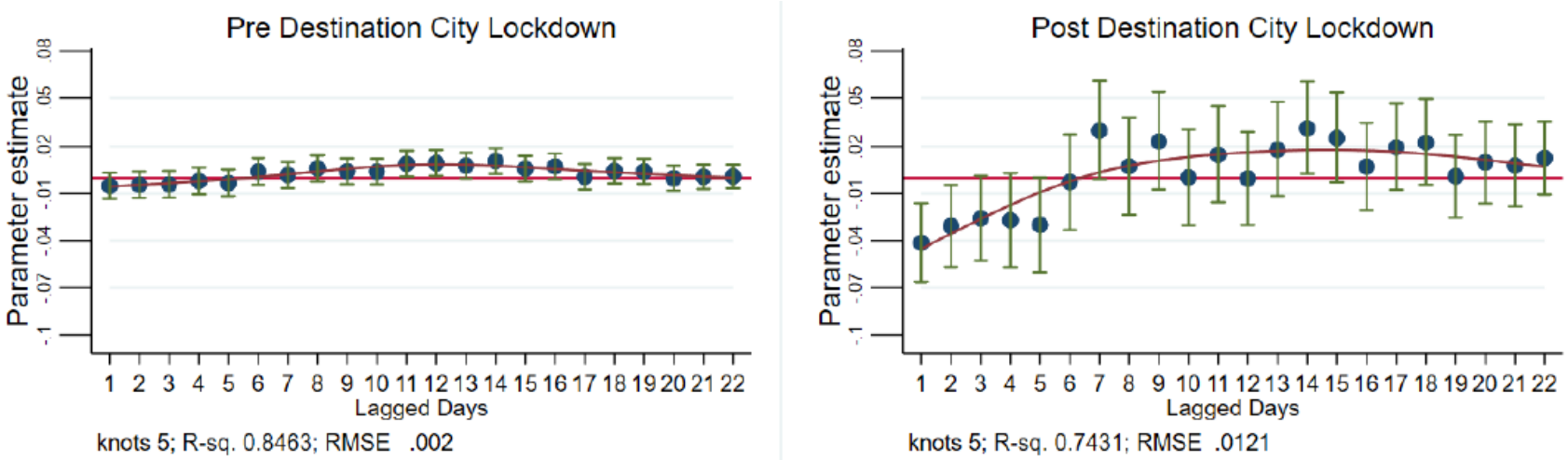


(a) Lagged Effects of Wuhan Inflows: Pre and Post Destination Cities' Lockdown

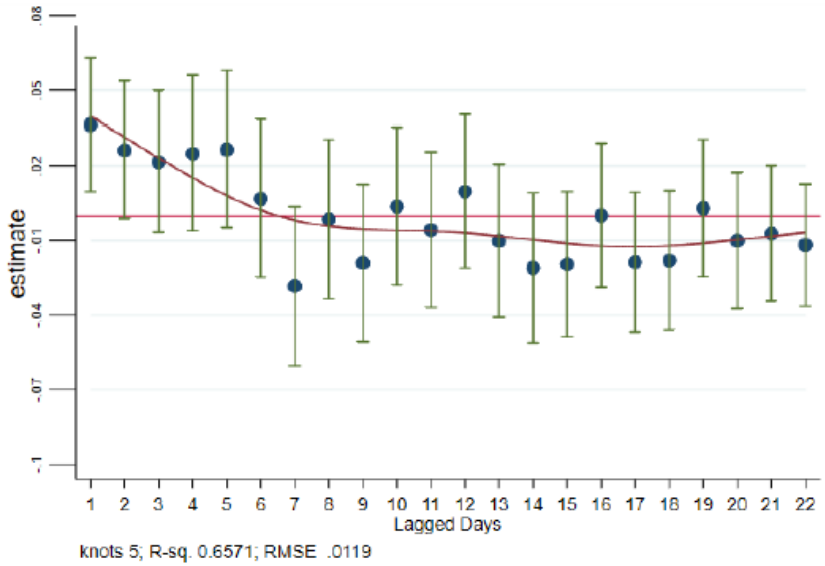


(b) Difference in the Effects

Results: Effect of Social Distancing on Hubei Inflow



(a) Lagged Effects of Wuhan Inflows: Pre and Post Destination Cities' Lockdown



(b) Difference in the Effects

Conclusions

- The lockdown of Wuhan reduced inflow into Wuhan by **76.64%**, outflows from Wuhan by **56.35%**, and within-Wuhan movements by **54.15%**
- The largest impact on the newly confirm cases today comes from the inflow population from Wuhan or other cities in Hubei about **12 to 14 days earlier**
- The number of officially reported cases in Wuhan was **42.17%** lower than our estimate on Jan 23, 2020, and **11.33%** lower as of Feb 29, 2020
- In the absence of Wuhan lockdown, the COVID-19 cases would be **64.81% higher** in the 347 Chinese cities outside Hubei province, and **52.64% higher** in 16 non-Wuhan cities inside Hubei
- Social distancing policies are **effective** in reducing the spread of 2019-nCoV virus in the destination cities