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

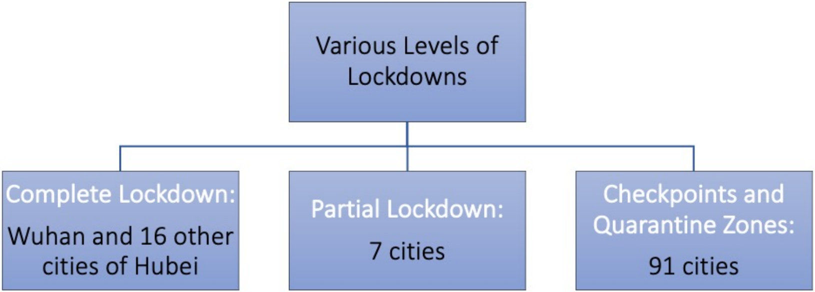
[3]The Chinese University of Hong Kong

AEA/ASSA 2021 Virtual Annual Meeting
Jan 3, 2021

Key Known Information of 2019-nCoV

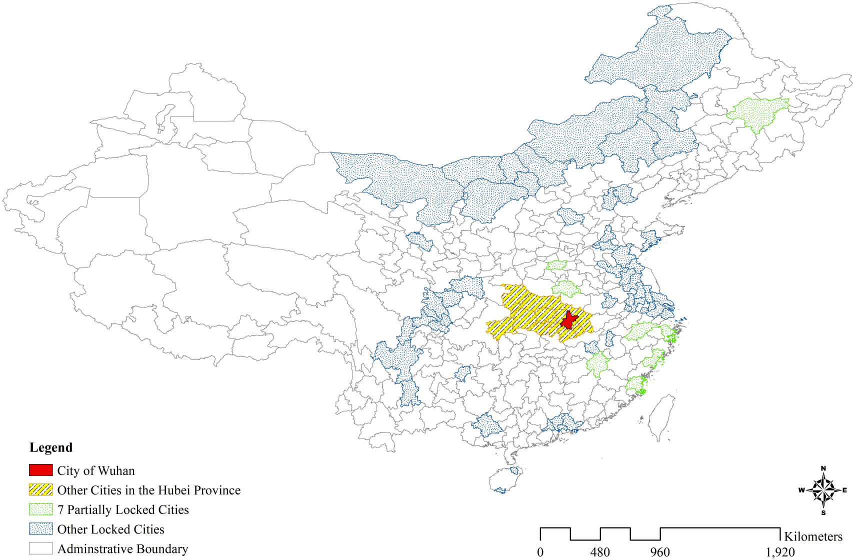
- First detected in Wuhan, China, with the first case reported in early to mid December 2019 ("pneumonia with unknown origin")
- Common symptoms at onset of illness were fever, cough, and myalgia or fatigue (Huang et al., 2020, *Lancet*)
- China publicly confirmed human-to-human transmission on **Jan 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)

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

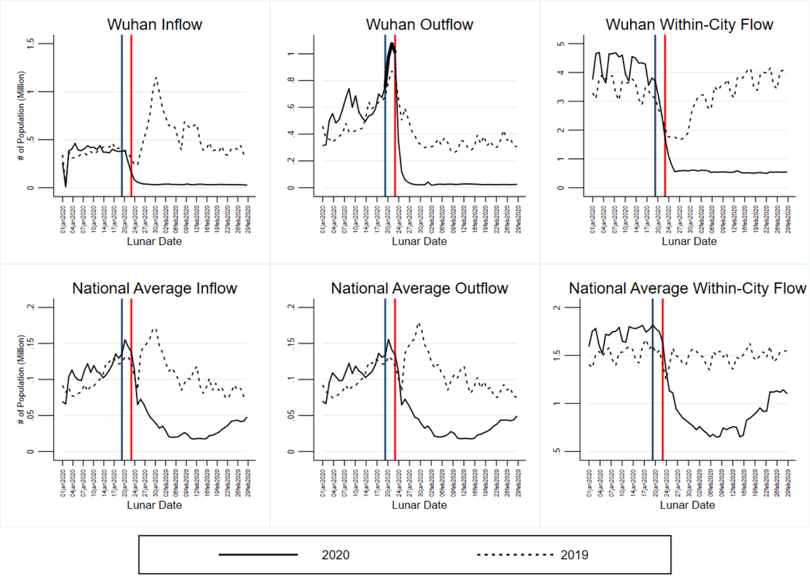
1. How does the lockdown of Wuhan affect population movement?
2. How do outflows from Wuhan and other cities in Hubei province affect virus infection in the destination cities?
3. Are social distancing policies in the destination cities effective in reducing the spread of the infections?
4. How many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

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, 2020
    - 90,848 person movements per inter-city index unit
    - 2,182,264 person movements per within-city index unit
  - 2,977,899 city-pair-day observations for 120,142 pairs of cities for 364 Chinese cities
  - 21,840 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:

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

  - $\mathit{Ln(WithinCityFlow}_{i,t})$ is the logarithmic within-city population mobility measure for city i at date t
  - $\mathit{Before}_{1,t}$, $\mathit{Before}_{2,t}$ 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 Inflow

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
Fixed Effects	City-Pair FE and Daily FE in Panel A and Panel B City FE and Daily FE in Panel C	

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

Results: Impact of Lockdown on Outflow

Treatment Group	Wuhan 2020	Wuhan 2020
Control Group	Wuhan 2019	7 Cities 2020
Treatment Effects	Lockdown + Virus	Lockdown
Model	(1)	(2)
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

Fixed Effects	City-Pair FE and Daily FE in Panel A and Panel B City FE and Daily FE in Panel C

Results: Impact of Lockdown on Within-city Flow

Treatment Group	Wuhan 2020	Wuhan 2020
Control Group	Wuhan 2019	7 Cities 2020
Treatment Effects	Lockdown + Virus	Lockdown
Model	(1)	(2)
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	

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

- $Treat * Before_2$ from Model 1: Panic Effect
- $Treat * After$ from Model 1: Lockdown+Virus Effect
- $Treat * After$ from Model 2: Lockdown Effect

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%***

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:

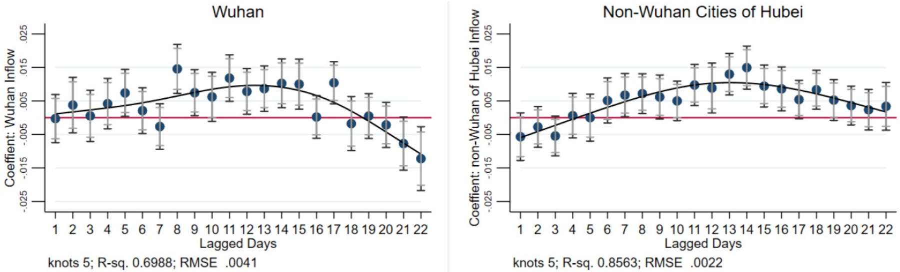
$$\begin{aligned} Ln(1{+}NewCase_{i,t}) &= \alpha + \sum_{\kappa=1}^{22} \beta_{1\kappa} \cdot Ln\left(Inflow_{i,WH,t-\kappa} \cdot I_{WH,t-\kappa}\right) \\ &+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in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

  - 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, 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.
  - City fixed effects μ_i and date fixed effects θ_t are included
  - The standard errors are clustered at the date level

Data

- 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
  - assume a constant ratio of asymptomatic and symptomatic cases

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



- 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

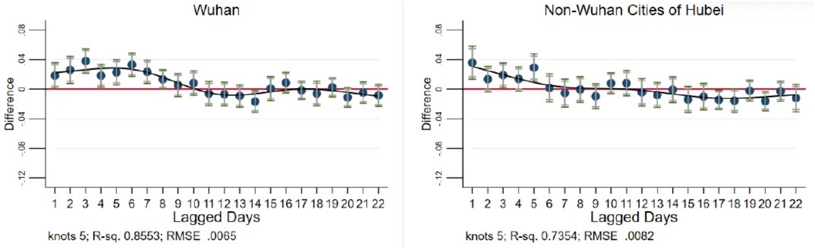
Question 3: What is the effect of social distancing on virus transmission?

- Specification:

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

  - $Lockdown_{i,t} = 1$ if time t is a date after destination city i's "lockdown" date, and 0, otherwise.

Results: Effect of Social Distancing



- 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

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:

$$
\begin{aligned}
Ln(1+\widetilde{NewCase}_{i,t}) = \hat{\alpha} + \sum_{\kappa=1}^{22} \hat{\beta}_{1\kappa} \cdot Ln\left(\widetilde{Inflow_{i,WH,t-\kappa}} \cdot I_{WH,t-\kappa}\right) \\
+ \sum_{\kappa=1}^{22} \beta_{2\kappa} \cdot Ln\left(\sum_{j\neq i, j\neq WH, j\in HB} Inflow_{i,j,t-\kappa} \cdot \sum_{j\neq i, j\neq WH, j\in HB} I_{j,t-\kappa}\right) + \hat{\mu_i} + \hat{\theta_t} + \hat{\epsilon_{i,t}},
\end{aligned}
\tag{5}
$$

  - $\hat{\beta}_{1\kappa}$ 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 $\hat{\mu_i}$ and date fixed effects $\hat{\theta_t}$ 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	-11.49%	$\boxed{106.06\%^{***}}$	-24.04%***
Virus Effect	-64.97%***	$\boxed{-37.31\%^{***}}$	-64.73%***
Lockdown Effect	-76.98%***	-56.31%***	-55.91%***

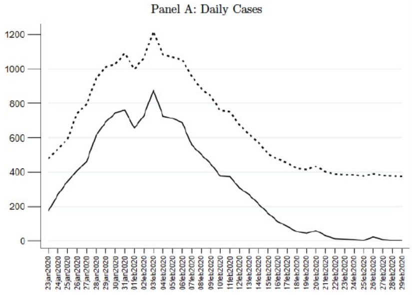
- In the absence of Wuhan lockdown, we would expect that outflows from Wuhan in days after Jan 23 to be:

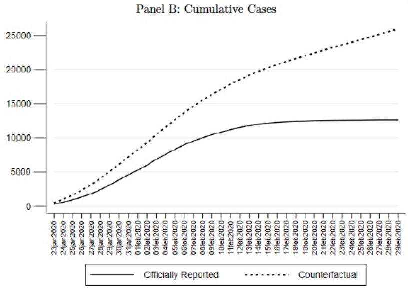
$$\overbrace{(1-0.3731)}^{\text{Reduction from Virus Effect}} \quad * \quad \overbrace{(1+1.0606)}^{\text{Increase from Panic Effect}} \quad = 1.29 \tag{6}$$

  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





Conclusions

- The lockdown of Wuhan reduced inflow into Wuhan by **76.98%**, outflows from Wuhan by **56.31%**, and within-Wuhan movements by **55.91%**
- 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**
- Social distancing policies are **effective** in reducing the spread of 2019-nCoV virus in the destination cities
- In the absence of Wuhan lockdown, the COVID-19 cases would be **105.27% higher** in the 347 Chinese cities outside Hubei province as of Feb 29, 2020