Population migration, spread of COVID-19, and epidemic prevention and control: empirical evidence from China

Zhen Hu1*, Yuanyang Wu2, Hualei Yang2, Lin Xie3, Anqi Zhang4, Xueyu Lin2

1 School of Economics and Management, Northwest Agricultural and Forestry University
2 School of Public Administration, Zhongnan University of Economics and Law
3 Institution of Population and Labor Economics, The Chinese Academy of Social Science
4 School of Public Economy and Management, Shanghai University of Finance and Economics

* Corresponding Author:

Zhen Hu
huzhenjingji@163.com
Tel: +8617782872897
Abstract

Background: This study applied the SEIR model to analyze and simulate the transmission mechanisms of the coronavirus disease 2019 (COVID-19) in China.

Methods: The population migration was embedded in the SEIR model to simulate and analyze the effects of the amount of population inflow on the number of confirmed cases. Based on numerical simulations, this study used statistical data for the empirical validation of its theoretical deductions and discussed how to improve the effectiveness of epidemic prevention and control considering population migration variables. Statistics regarding the numbers of infected people in various provinces were obtained from the epidemic-related data reported by China’s National Health Commission.

Results: This study explored how the epidemic should be prevented and controlled from the perspective of population migration variables. It found that a combination of a susceptible population, an infected population, and transmission media was an important route affecting the number of infections and that the migration of a Hubei-related infected population played a key role in promoting epidemic spread. Epidemic prevention and control should focus on regions with better economic conditions than the epidemic region. Prevention and control efforts should focus on the more populated neighboring provinces having convenient transportation links with the epidemic region. To prevent and control epidemic spread, priority should be given to elucidating the destinations and directions of population migration from the domestic origin of infections, and then stemming population migration or human-to-human contact after such migration.

Conclusions: This study enriched and expanded on simulations of the effects of population migration on the COVID-19 epidemic and China-based empirical studies while offering an
epidemic evaluation and warning mechanism to prevent and control similar public health emergencies in the future.

**Keywords:** Wuhan, COVID-19, population migration, epidemic prevention and control
Background

In December 2019, patients with pneumonia of unknown etiology were reported as having been in the Huanan Seafood Wholesale Market in Wuhan, Hubei province. They were admitted to the Wuhan Jinyintan Hospital for treatment and were later diagnosed with coronavirus disease 2019 (COVID-19). Since this incident, COVID-19 has been brought into the public eye [1–4]. With the large-scale human migration that took place during the Chinese New Year, COVID-19 spread rapidly in China. As of May 2, 2020, there had been 84,391 confirmed cases and 4,643 deaths in China, with a mortality rate of 5.5%. Among them, 50,332 cases were reported in Wuhan, where the mortality rate was as high as 7.69%. Based on epidemiological theories, the combination of a susceptible population, an infected population, and transmission media is an important channel that affects the number of infections. In particular, the migration of exposed and symptomatic infected people during the incubation period play a key role in promoting the spread of the epidemic. In order to stem population migration, various Chinese provinces and cities successively activated Level-1 Response to Major Public Health Emergencies, carried out joint prevention and control measures, work and production stoppages, as well as enacting stringent lockdowns in all urban and rural communities [5–7]. These measures eventually succeeded in keeping China’s epidemic at bay.

While China was seeing the first signs of success in its epidemic control efforts, countries around the world succumbed to the disease one by one due to international population migration in the context of globalization. As of May 2, 2020, 192 countries and regions around the world had reported COVID-19 cases, with the cumulative confirmed cases totaling 3.33 million and a death
toll exceeding 245,400 worldwide. The mortality rate is up to 7.3% globally and it is a staggering
14% in countries like the UK and France. Even though countries have acted against COVID-19, the results have fallen short [8]. In the face of this global public health crisis, Bruce Aylward, senior advisor to the Director-General of the World Health Organization (WHO), called for countries to strengthen their collaboration and management in aspects such as scientific research, population movement, and the supply of medical resources and to add greater synergy to the global battle against the challenges posed by COVID-19. In view of the initial success obtained by China in COVID-19 prevention and control, Tedros Adhanom, WHO’s Director-General, urged countries around the world to learn from China’s experience [9]. In light of China’s anti-epidemic measures, the organization and mobilization capabilities of the government, manifest especially in the lockdowns and other measures intended to stem population migration, have attracted widespread attention from other countries [3, 10–13]. To contribute China’s experience to the global fight against COVID-19 and to respond to similar public health emergencies in the future, there is a need for China to summarize the patterns of disease progression and anti-epidemic experience obtained from this event [14, 15], with an emphasis on the mechanism through which population migration affects epidemic spread.

In conclusion, contact between the infected population and a susceptible population is an important channel for disease transmission, in which migration plays a key role in promoting their combination. In this study, the population migration perspective was used in an attempt to answer two questions: 1) How did population migration affect the number of confirmed cases in each region and what were the transmission patterns of the epidemic? 2) Based on theoretical and empirical analyses of the effects of population migration on the number of confirmed cases, what
were the factors influencing population migration, and how should a warning system be constructed for the prevention and control of the current epidemic and any similar event in the future? The specific content of this paper is as follows: Firstly, this paper applied the SEIR model to simulate the effects of the amount of Hubei-related population inflow on the number of infections in various Chinese provinces, followed by an analysis of the factors influencing population migration. Secondly, this paper drew on population migration data and statistics of confirmed cases to empirically prove the theoretical deductions in the previous section. Thirdly, this paper elaborated on how to construct an epidemic evaluation and warning system based on population migration variables.

This study was distinctive from the studies by Wu et al. [16], Zhao and Chen [17], Tang et al. [18], Anastassopoulou et al. [19], Yang and Wang [20], Chen et al. [21] and Huang et al. [22] in the following ways: First, population migration was embedded in the SEIR model to simulate and analyze the effects of the amount of population inflow on the number of confirmed cases. Second, compared with existing studies based on numerical simulations, this study used statistical data for the empirical validation of its theoretical deductions. Finally, compared with existing numerical simulations and empirical studies along the same lines, this study also discussed how to improve the effectiveness of epidemic prevention and control based on population migration variables. The contributions of this study lay in enriching and expanding on simulations of the effects of population migration on the COVID-19 epidemic and China-based empirical studies. At the same time, it provided an epidemic evaluation and warning mechanism based on population migration variables to prevent and control similar public health emergencies in the future. In addition, it summarized China’s epidemic prevention and control measures based on population migration
variables to assist in the global efforts in disease prevention and control that are underway.

**Theoretical analysis**

**Construction of SEIR model**

The studies by Chen et al. [21], Yang and Wang [20], Anastassopoulou et al. [19], Tang et al. [18], Zhao and Chen [17], Wu et al. [16], Huang et al. [22], and Wan et al. [23] were used as references for the SEIR model settings in the present study. The model settings were as follows:

Considering that COVID-19 had an incubation period and assuming that populations were divided into five compartments, $S_t$, $E_t$, $I_t$, $R_t$, and $D_t$ denoted the sizes of the susceptible population, population in the incubation state, infected population, recovered population, and deceased population at time $t$ respectively. The total number of people in the system, $N$, was a constant and assumed to be $\sum_{i=S,E,I,R,D} t(t) = N$.

Assuming that individuals in the incubation and infected states came into contact with $m_1$ and $m_2$ people on average daily respectively, that the levels of transmissibility in these two groups were $\beta_1$ and $\beta_2$ respectively, and that the daily numbers of people infected by individuals in the incubation and infected states were $m_1 \beta_1 ES / N$ and $m_2 \beta_2 IS / N$ respectively, then the rate of change in the size of the susceptible population was defined as follows:

$$\frac{dS}{dt} = -m_1 \beta_1 ES / N - m_2 \beta_2 IS / N \quad (1)$$

Assuming that individuals in the incubation state converted to infected individuals every day at a probability of $\alpha$, then the rate of change in the size of population in the incubation state was defined as follows:
\[
\frac{dE}{dt} = m_1 \beta_1 ES / N + m_2 \beta_2 IS / N - \alpha E
\]  
(2)

Assuming that infected individuals converted to the recovered and deceased states every day at a probability of \( \gamma_1 \) and \( \gamma_2 \) respectively, then the rates of changes in the sizes of the infected population, recovered population, and deceased population were defined as follows:

\[
\frac{dI}{dt} = \alpha E - \gamma_1 I - \gamma_2 I
\]  
(3)

\[
\frac{dR}{dt} = \gamma_1 I
\]  
(4)

\[
\frac{dD}{dt} = \gamma_2 I
\]  
(5)

In summary, the SEIR model was defined as follows:

\[
\begin{align*}
\frac{dS}{dt} &= -m_1 \beta_1 ES / N - m_2 \beta_2 IS / N \\
\frac{dE}{dt} &= m_1 \beta_1 ES / N + m_2 \beta_2 IS / N - \alpha E \\
\frac{dI}{dt} &= \alpha E - \gamma_1 I - \gamma_2 I \\
\frac{dR}{dt} &= \gamma_1 I \\
\frac{dD}{dt} &= \gamma_2 I
\end{align*}
\]  
(6)

Initial parameter settings

According to Geng et al. [24], \( \beta_1 = \beta_2 = 0.045 \). According to Fan et al. [25], \( \alpha = 1 / 7 \).

Based on the reported COVID-19 data by China’s National Health Commission, data was publicly available, the proportion of recovered patients and the mortality rate fluctuated at 0.049-0.085 and 0.0009-0.0015 respectively between 1 and 6 March 2020, and the mean values of the intervals
were obtained so that $\gamma_1 = 0.069$ and $\gamma_2 = 0.00115$. Assuming that $I_t, R_t$, and $D_t$ were all 0 initially, $N = E_0 + S_0 = 1$. Considering the complexity of the solution process, the results of the numerical simulation are shown directly.

*Theoretical analysis of the effects of population migration*

After the onset of the outbreak, two main types of people migrated from Hubei to other provinces. They were permanent Hubei residents from other provinces who returned home during the Chinese New Year and permanent local residents of Hubei province who travelled to visit relatives or as outbound tourists during the Chinese New Year. Both groups migrated to other provinces after the COVID-19 outbreak occurred. The number of potential infections during such migration mainly affected the initial number of people in the incubation state in each city, $E_0$.

Hence, an analysis of the effects of the Hubei-related population migration on the number of infections in other provinces resolved itself into an analysis of the effects of the initial number of people in the incubation state, $E_0$, on the number of infected people, $I(t)$. Considering that the number of infected people was dynamic and that the trajectory of the number of newly infected people displayed an inverted $U$ shape\(^1\), only the effects of the initial number of people in the incubation state on the peak number of infected people were obtained. In order to ensure the robustness of the conclusion, simulation studies were carried out under two states: frequent public interaction and infrequent public interaction. When public interaction was frequent, $m_1 = 5$ and $m_2 = 3$. When public interaction was infrequent, $m_1 = 3$ and $m_2 = 1$. The simulation results

\(^1\) See the appendix for the inverted U shape.
are shown in Figure 1, where the y-axis is the peak number of infected people, and the x-axis is the proportion of Hubei-related population inflow in the total population of a given province.

As shown in Figure 1, under both frequent and infrequent public interaction, a greater number of people migrating from Hubei to other provinces, i.e., a greater initial number of people in the incubation state, was associated with a higher peak number of infected people. The peak number of infected people was higher when there was frequent public interaction, compared to when there was less frequent interaction. In conclusion, the theoretical analysis revealed that the greater the initial number of people who migrated from Hubei after the outbreak, the greater the number of infected people in other provinces at a later stage. To validate the aforementioned proposition from the empirical angle, the next section analyzed the empirical relationship between the number of people who migrated from Hubei to other provinces, or the number of those who migrated from other provinces to Hubei, and the number of infected people in those provinces.

The above theoretical analysis revealed that contact between the infected and susceptible populations was the main cause of epidemic spread. In the context of China’s large-scale population migration during the Chinese New Year, migration of infected people from Hubei played a key role in determining the speed of epidemic spread. The answers to questions like whether these theoretical deductions were empirically valid, how population migration influenced the number of infected people in various Chinese provinces, and which variables were influencing population migration would help to prevent and control the epidemic at hand and establish warning indicators for similar epidemics in the future. In the next section, empirical validation was conducted to substantiate the theoretical deductions made above.
Methodology Model settings

Based on the theoretical analysis, the number of infected people was mainly affected by the combined effects of factors including the number of susceptible people, the number of infected people, transmission media, medical conditions, and immunization measures. The number of infected people in another province (\( Y \)) was seen as the dependent variable. The number of susceptible people was equated roughly to the number of permanent residents in each province. Considering that the initial number of infected people was associated with Hubei, the size of a Hubei-related population could be equated roughly to the number of infected people carrying the virus. Thus, the number of permanent residents in each province (\( P \)) were taken as the number of susceptible people. Considering that the outbreak site of COVID-19 was in Hubei Province, the number of people who migrated from Hubei to each province (\( PMH \)) and the number of people who migrated from each province to Hubei (\( PHM \)) were taken as the major source of infection in each province. Population density (\( PD \)) and traffic density (\( TD \)) were considered to be the transmission media for the epidemic. Also, the proportion of population over 65 years old (\( Over65 \)), and whether to start the first level response before January 24 (\( Emergency \)) were taken as the control variables in the model. This generated the susceptible population, the infected population, and transmission media in the theoretical model mentioned above.

In addition, this study also shed light on factors affecting population migration from Hubei to other provinces and vice versa. By drawing on the studies by Karemera et al. [26], Piore [27], Stark and Taylor [28], Cai and Wang [29], Shen [30], Henry et al. [31], Fan [32], Zhu and Chen [33], and Borrow [34] on the factors influencing population migration, the following variables were selected as the control variables: the number of permanent residents in each province, whether the province neighbors Hubei (\( nei ghbor \)), whether the province is in Southern China (\( regi on \)), the number of high-speed trains between the province and Hubei (\( QHai l \)), the geographical distance between the province and Hubei (\( di st an \)), the comparison of disposable income per capita between the province and Hubei (\( compari son \)) and the comparison of
The number of 5A-level scenic spots between the province and Hubei \( (\text{FiveA}) \). The definition of each variable is shown in Table 1. The number of infected people and the model of population migration were defined as follows:

\[
Y_i = \alpha + \beta_i X_i + \mu_i Z_i
\]  
(9)

\[
M_i = c + \mu_i D_i + \theta_i
\]  
(10)

where \( i \) represents the \( i \)th province; \( Y_i \) is the main explanatory variable; \( M_i \) is the explained variable of the model (10); \( X_i \) is the main explanatory variable of the model (9); \( Z_i \) is the control variable of the model (9); and \( D_i \) denotes the explanatory variable of model (10). \( \beta_i \), \( \gamma_i \), and \( \mu_i \) represent the parameters to be estimated for models (9) and (10). \( \epsilon_i \) and \( \theta_i \) denote the random error terms, namely, other factors affecting the number of infected people.

**Descriptive statistics**

Statistics regarding the numbers of infected people in various provinces were obtained from the epidemic-related data reported by China’s National Health Commission. The number of people who migrated from Hubei to other provinces and vice versa was obtained from the 2015 China 1% National Population Sample Survey [35]. The number of permanent residents, transport mileage, and disposable income per capita for each province, the proportion of population over 65 years, and the number of 5A-level scenic spots were sourced from the 2015 China Statistical Yearbook [36]. The surface areas of various provinces were obtained from the Ministry of Land and...
The Qinling–Huaihe Line was used as a reference line to distinguish between Northern and Southern China, with provinces to the north of this line as northern provinces and those to the south as southern provinces. In terms of the number of inter-provincial high-speed trains, only the number of high-speed trains operating between each provincial capital and Wuhan was considered. The official website of the China Railway Corporation (12036.cn) was used to obtain the number of high-speed trains between each province and Wuhan. As for the other variables, the population density was obtained by dividing the number of permanent residents in each province by the land area. Traffic density was equivalent to the total mileage of roads, railways, and waterways in each province divided by the land area. Whether the province has initiated the first level response to the public health emergencies was obtained from the local government website. The results of the descriptive statistics of various variables are detailed in Table 1. The number of confirmed cases in each province is shown in Figure 2.

Table 1 shows that as of 2 May 2020, the mean number of infected people in various provinces was 2,619. In provinces outside of Hubei, the number of infected people was the highest in Guangdong (n=1,395) and lowest in Tibet (n=1). The mean number of people who migrated from Hubei to another province was 318,700; the greatest number of people migrated to Guangdong and the least to Tibet. The mean number of people who migrated from another province to Hubei was 98,910; the greatest number of people was from Henan (n=448,900) and the least was from Ningxia (n=6,000). The mean number of permanent residents in each province was 43,950,320 people; the greatest number of people was in Guangdong and the least was in Tibet. The province with the highest population density was Shanghai (3,825 people per km²).
and the one with the lowest population density was Tibet (2 people per km$^2$). Shanghai had the highest traffic density, while Tibet had the lowest.

Population migration and number of infected people

To observe the relationship between population migration and the number of infected people in each province, the numbers of people who migrated from Hubei to other provinces and vice versa as well as the numbers of permanent residents and infected people in various provinces were tabulated in Table 2. As shown in the second column, the provinces with the highest number of migrants from Hubei, in descending order, were Guangdong, Zhejiang, Shanghai, Jiangsu, Beijing, and Fujian, all of which are more economically developed than Hubei. Meanwhile, the provinces with the least number of migrants from Hubei, in ascending order, were Tibet, Ningxia, Inner Mongolia as well as Liaoning, Jilin, and Heilongjiang (collectively termed the Northeastern China), all of which are the less economically developed regions in the northwest and northeast China. According to the fourth column, the provinces with the greatest number of people who migrated to Hubei were largely from the peripheral provinces of Hubei. They were, in descending order, Henan, Hunan, Guangdong, Chongqing, Anhui, and Jiangxi. The provinces with the least number of people who migrated to Hubei, in ascending order, were Qinghai, Tibet, Ningxia, Jilin, Liaoning, and Tianjin, which are in northwest and northeast China. In contrast, the provinces with the most permanent residents, in descending order, were Guangdong, Shandong, Henan, Sichuan, Jiangsu, and Hebei. Those with the least permanent residents, in ascending order, were Xinjiang, Qinghai, Tibet, Ningxia, and Hainan.

Outside of Hubei, the provinces with the highest number of infected people, in descending
order, were Guangdong, Zhejiang, Henan, Hunan, Anhui, and Jiangxi. They could be divided into two categories. The first was the major destinations for population migration from Hubei, including Guangdong, Zhejiang, and other economically developed provinces. The other category was the major origins of population migration to Hubei, including peripheral provinces such as Henan, Hunan, Anhui, and Jiangxi. The least infected provinces were mostly places in northwest and northeast China that were unpopular among Hubei residents, such as Tibet, Qinghai, Xinjiang, Ningxia, Inner Mongolia, and Jilin. This intuitively proved that population migration has a profound effect on the number of infected people in each province, with the major destinations for migration from Hubei and major sources of migration migrating to Hubei as the hardest hit regions. In contrast, northeast and northwest provinces seldom visited by Hubei residents and from which fewer people migrated to Hubei were less affected by the epidemic. To prove the above viewpoint with greater rigor, a quantitative analysis was carried out in the following section.

Results are shown in Table 2.

Please place Table 2 here.

In this paper, the empirical strategy of this paper is as follows: on the one hand, this study analyzed the impact of population migration factors on the number of infected people. The outcome variable is the number of infected people and the explanatory variables mainly include $P$, $PHM$, $PMH$, $PD$, $TD$, $Over65$ and $Emergency$; on the other hand, this study further analyzed the factors affecting population migration. The outcome variable is the number of population migration, and the explanatory variables include $P$, $neighbor$, $region$, $QHrail$, $distan$, $comparison$ and $FiveA$.

Results
Analysis of migration factors affecting infected people

Contact between the susceptible population and the infected population is the main channel of infection in which population migration plays a critical role. In order to empirically validate the above deduction, four OLS models were generated (Table 3). In the four models, the dependent variables are the number of infected people. In Model 1, the number of permanent residents in each province was added as independent variable. In Model 2, the number of people who migrated from Hubei to other provinces and the number of people who migrated from other provinces to Hubei were added as independent variables. In Model 3, the interaction terms added were as follows: (1) between the number of permanent residents in each province and the number of people who migrated from Hubei to that province, and (2) between the number of permanent residents in each province and the number of people who migrated from that province to Hubei were added. Model 4 was an extension of Model 3, with the addition of variables representing the transmission media, demographic characteristics and response strategies from provinces. Considering that the frequency of contact between the susceptible population and infected population was hardly a measurable condition for human contact, population density and traffic density were used to measure the transmissibility associated with the transmission media. In this study, the focus of the discussion was placed on the regression results in Model 4.

Please place Table 3 here.

From Models 1–4, the following findings could be derived: Firstly, Model 1 showed that the permanent population coefficient was positive but insignificant, which suggested that the number
of infected people would be zero in provinces without a source of infection. Model 2 showed that
the greater the number of people migrating from Hubei to another province or vice versa, the
greater the number of infected people in that province. Considering that the outbreak site of
COVID-19 was in Hubei Province, the explanation was two-fold in the context of population
migration during the Chinese New Year: while permanent Hubei residents from other provinces
who were infected returned home during the holiday, infected local residents migrated from Hubei
to other provinces for tourism, business, and visiting purposes. This increased the sources of
infection in the corresponding provinces. Due to contact between the sources of infection and the
susceptible population, every increase of 10,000 in the number of people migrating from Hubei to
each province and vice versa would raise the number of infected people in that province by 1.953
and 30.89, respectively.

According to Model 3, the interaction terms between the number of permanent residents in
each province and the number of people who migrated from Hubei to that province, and between
the same and the number of people who migrated from that province to Hubei, significantly and
positively affected the number of infected people. This validated the theoretical deduction made
above that the number of infected people depends on the contact and intensity of contact between
the susceptible population and the infected population. In other words, the number of infected
people is directly proportional to the product of the number of susceptible people and infectious
agents. Considering that the infectious agents are dependent on the transmission media, Model 4
showed that population and traffic density had significant effects on the number of infected people
in each province, while the population density, proportion of population over 65 years old and
whether to start the first level response before January 24 had insignificant effects on the number
of susceptible people in each province. The greater the traffic density, the more convenient the traffic, which is conducive to population migration and will accelerate the spread of the epidemic. Therefore, strengthening traffic control plays an important role in controlling the spread of the epidemic.

Analysis of factors affecting population migration

From the above analysis, it could be concluded that the key factor affecting the number of infected people in each province is the regulation of contact between permanent residents and the infected population in Hubei, particularly those migrating from Hubei to other provinces or vice versa. Hence, the population migration during the Chinese New Year holiday served as a key determinant affecting the number of infected people. On the other hand, studying the factors affecting the number of people migrating from Hubei to various provinces and vice versa will help to determine the effects of population migration-related factors on the spread, prevention, and control of the epidemic. This is particularly beneficial to the global efforts in pandemic prevention and control at play as well as the establishment of an epidemic warning system. Thus, the number of people who migrated from Hubei to each province, the number of people who migrated from each province to Hubei, and the sum of the two were selected as the explained variables to generate Models 5–7. OLS regression analysis was then conducted, the results of which are shown in Table 4.

From Table 4, the following conclusions could be drawn. Firstly, economic condition is an
important factor affecting outward migration from Hubei. A larger number of Hubei residents migrated to more economically developed provinces. Compared with provinces with lower disposable income per capita than Hubei, provinces with higher disposable income per capita attracted, on average, an additional 454,500 people as visitors. Secondly, people from neighboring provinces were more likely to migrate to Hubei. Compared with non-neighboring provinces, they saw, on average, 84,640 more people migrating outwardly to Hubei. A greater number of high-speed trains to and from Hubei was also associated with a greater number of people migrating to Hubei. Every increase of one in the number of high-speed trains would increase the number of people migrating to Hubei by 1,380 people. People from provinces with more 5A-level scenic spots were also more willing to migrate to Hubei, which may be related to the rich tourism resources in Hubei. Thirdly, the number of migrants related to Hubei Province is only affected by the relative economic level with Hubei Province.

Please place Table 4 here.

Discussion

The above theoretical and empirical analyses identified local Hubei residents who migrated to other provinces and residents from other provinces who migrated to Hubei as key enablers of the spread of the epidemic. One of the sources of infection was the population migrating out of Hubei. Economic factors played a crucial role in affecting the outward migration of the Hubei population. Thus, from the perspective of population migration variables, provinces with higher disposable income per capita than Hubei should be regarded as the key regions for epidemic prevention and control, as they were among the top-ranked regions with the most population
inflows from Hubei. For example, Guangdong and Zhejiang, both of which had higher disposable income per capita than Hubei, were the first and second most preferred destinations for outward migration by Hubei residents. Large-scale outward migration increases the contact between the susceptible population and infected population, which accelerates the spread of the epidemic. Statistics indicate that, as of 2 May 2020, Guangdong and Zhejiang ranked first and third respectively among provinces outside of Hubei in terms of the number of confirmed cases excluding imported cases. After the outbreak of the epidemic, both of these provinces recognized the problem with acute awareness and became the first Chinese provinces to activate the Level-1 Response to Major Public Health Emergencies. They took proactive steps to launch a series of epidemic prevention and control policies, which effectively stalled the spread of the epidemic.

Population migration from other provinces to Hubei constitutes yet another major source of infection. The above theoretical and empirical analyses found that from the perspective of population migration, provinces neighboring Hubei and those that are linked with Hubei via a convenient transport network saw a larger number of people migrating to Hubei. These people were likely to return home during the Chinese New Year, which added to the number of infected people in their home provinces. Hence, epidemic prevention and control efforts should be targeted at these provinces. Henan and Hunan provinces serve as examples. Both provinces are geographically adjacent to Hubei. In terms of transportation, there are frequent high-speed trains operating between both provinces and Hubei, which makes transportation highly convenient. During the Chinese New Year, a large number of people were recorded migrating from Hubei to both provinces. It is noteworthy that since the outbreak of COVID-19, Henan and Hunan have both been severely hit by the epidemic. As of 2 May 2020, the numbers of confirmed cases in
Henan and Hunan were the second and fourth highest among Chinese provinces. From the onset of the outbreak, the Henan and Hunan provincial governments had been acutely aware of this problem, with the former being the more perspicacious of the two. The Department of Public Security of Henan Province immediately established a steering committee for epidemic prevention and control, and the public healthcare sector went into a system-wide state of emergency. At the grassroots level, stringent lines of defense were first set up to minimize the likelihood of any contact between the susceptible population and sources of infection to prevent the rapid spread of the epidemic.

In summary, based on theoretical and empirical analyses of the effects of population migration on the number of confirmed cases in China, in the future China should adopt the following measures for effectively preventing and controlling the epidemic spread: Efforts should first be made to specify the directions of outward migration from the domestic origin of infections and define the key regions for epidemic prevention and control based on demographic variables. After locating the key regions, contact between the susceptible population and infected population should be effectively cut off within these regions. The successful outcomes of China’s epidemic prevention and control efforts have attested to the effectiveness of this strategy.

Implication

This study first applied the SEIR model to analyze and simulate transmission mechanisms of COVID-19 in China to identify transmission mechanism. In addition, this paper analyzed the
migration factors affecting infected people and further analyzed the factors affecting population migration. The conclusions are as follows: given that the transmission media and medical conditions remained unchanged, the key mechanism underlying the transmission of COVID-19 was human contact between the susceptible population and the infected population, promoted mainly via Hubei-related population migration; empirical analysis found that population migration and traffic density have a significant positive impact on the number of infected people in China; Further analysis found that the relative economic development level and traffic convenience have an important impact on population migration.

The policy implications of this paper are: when preventing and controlling the epidemic, it is necessary to have a thorough understanding of the population migration situation of the domestic origin of infections, clarify the direction of population migration and restrict population migration. They specifically include the following two aspects: On the one hand, countries and regions should determine the directions of population migration to and from the epidemic region, and identify whether a region is one of the main destinations or origins of population migration. For example, considering that economically developed areas such as Guangdong and Zhejiang are the preferred destinations for Hubei’s outbound migration, and the main sources of population migrating to Hubei are Henan and Hunan, which are adjacent to Hubei and have convenient transportation, so the epidemic prevention and control The above-mentioned provinces have maintained a high degree of vigilance and had effectively prevented and controlled the spread of the epidemic. For example, economically developed regions like Guangdong and Zhejiang were among the most preferred destinations for outward migration from Hubei, while populous and neighboring provinces such as Henan and Hunan, with convenient transportation links to Hubei,
were the main origins of migration to Hubei. Therefore, in terms of epidemic prevention and control, the above-mentioned provinces have maintained a heightened vigilance and have effectively prevented and controlled the spread of the epidemic.

On the other hand, there is a need to restrict population migration and reduce human-to-human contact during the epidemic, which requires the joint efforts of the cities, the communities, and the individuals. Combined with China’s experience, in terms of city, as Wuhan in Hubei Province was the outbreak site of the epidemic, to prevent the spread of the epidemic caused by population migration, a lockdown was declared, and all transportation links such as airports and railway stations were closed. Public buses, subways, ferries, and long-distance coaches temporarily ceased operations. The lockdown limited the spread of the epidemic to within the city. The shutdown of public transportation in the city minimized human-to-human contact and blocked the spread of the disease within the city. For cities other than Wuhan, firstly, temperature monitoring was performed at airports, seaports, railway stations, and bus stations, and people showing abnormal temperature would be immediately quarantined. Secondly, in-depth epidemiological surveys were conducted in cities to strengthen the tracking and management of close contacts. Information of people with travel or residence history in Wuhan or contact history with Wuhan residents was registered; Lastly, public places in the city were disinfected and gatherings were canceled. In terms of community, the body temperature monitoring and health registration carried out at the gates of the community can effectively screen suspected cases and high-risk infections. For residents’ food-related and daily necessities, community workers organized online order-taking for community residents, and community procurement was carried out before delivery to the home, thereby minimizing the risk of contact with infected people. In
terms of individuals, first, residents must comply with home quarantine and avoid going out to crowded places; second, residents should use online shopping channels and reduce contact with outside surrounding in daily life; Third, residents can also utilize online consultation to reduce the risk of infection in the process of consultation; finally, residents should utilize online work strategies to reduce the risk of contact with others at the workplace.

In addition, the limitations of this study are in the following two aspects: On the one hand, China’s population sample survey is conducted every five years. The data we use are from the 2015 China 1% Population Sample Survey Data. The latest census has not yet been conducted. Therefore, the latest data cannot be obtained; on the other hand, this paper is based on Chinese samples, and the conclusions obtained may have limitations and may not be applicable to other countries in the world.

List of abbreviations:

NCOV, number of infected people;
P, number of permanent residents in each province;
PD, population density;
TD, traffic density;
PHM, number of people who migrated from Hubei to another province;
PMH, number of people who migrated from another province to Hubei;
QHrail, number of high-speed trains between the provincial capital and Wuhan;
WHO, World Health Organization.

**Declarations**

**Ethics approval and consent to participate:**

Not applicable, not contain any individual persons data, ethics is not required as the data is publicly available.

**Consent for publication:**

Not applicable

**Availability of data and materials:**

The datasets generated and analyses during the current study are available in the National Bureau of Statistics of China and Chinese Center for Disease Control and Prevention, the data was publicly available.

**Competing interests:**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Funding:**

This work was supported by the Humanities and Social Sciences Fund of the Ministry of Education under (Grants number: 19YJC790167 & 18YJC790053); the China Postdoctoral Science Foundation (Grants number: 2018M631212), the Social Science Foundation of Shaanxi Province (Grants number: 2018S19).

**Author Contributions:**
H.Y. and Y. W. conceived of this research. Z.H. was responsible for the methodology. L.X. conducted software analyses. X.L. A. Z. and Z.H. conducted necessary validations. Y.W. conducted a formal analysis and handled the investigation. H.Y. gathered resources, curated all data, wrote/prepared the original draft, and was responsible for project administration. Z.H. and Y. N. reviewed and edited the manuscript, was responsible for visualization, supervised the project, and acquired funding. All authors participated in the critical revision of the manuscript and approved the final version.

Acknowledgements:

The authors thank Jianchao Luo and Rihong Zang for their constructive and valuable suggestions. The authors take responsibility for all remaining errors.
References

[1] Huang CL, Wang YM, Li XW, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, 2020, DOI: 10.1016/S0140-6736(20)30183-5.

[2] Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. New England Journal of Medicine, 2020, DOI: 10.1056/NEJMo2002032.

[3] Li Q, Guan XH, Wu P, et al. Early Transmission Dynamics in Wuhan, China of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine, 2020, DOI: 10.1056/NEJMoa2001316.

[4] Wang G, Jin X. The progress of 2019 novel coronavirus event in China. Journal of Medical Virology, 2020, DOI: 10.1002/jmv.25705.

[5] Chen T, Rui J, Wang Q, et al. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infectious Diseases of Poverty, 2020, 9(1): 24-24.

[6] Gong FH, Xiong Y, Xiao J, Lin L, Liu X D, Wang DZ, Li XK. China’s local governments are combating COVID-19 with unprecedented responses—from a Wenzhou governance perspective. Frontiers of Medicine, DOI: 10.1007/s11684-020-0755-z.

[7] Deng SQ, Peng HJ. Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China. Journal of Clinical Medicine, 2020, 9(1):575.

[8] Haffajee RL, Mello M. Thinking Globally, Acting Locally—The US Response to Covid-19. New England Journal of Medicine, 2020, DOI: 10.1056/NEJMp2006740.

[9] Prof Joseph TWuPhDa, KathyLeungPhDa, Prof Gabriel MLewungMDa. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet, 2020, 395(10225), 689-697.
[10] Kraemer MUG, Yang CH, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China [J]. Science, 2020, DOI: 10.1126/science.abb4218

[11] Prem K, Liu Yang, Russe T W, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study [J]. The Lancet, 2020, DOI: 10.1016/S2468-2667(20)30073-6.

[12] Kucharski AJ, Russe TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study [J], The Lancet, DOI: 10.1016/S1473-3099(20)30161-4.

[13] Tian HY, Liu YH, Li YD et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science, 2020, DOI: 10.1126/science.abb6105.

[14] Zhang H. Early lessons from the frontline of the 2019-nCoV outbreak [J]. The Lancet, 2020, 395(10225):687

[15] Xu JY, Chen YJ, Chen H, Cao B. 2019 novel coronavirus outbreak: a quiz or final exam [J]. Frontiers of Medicine, 2020, DOI: 10.1007/s11684-020-0753-1.

[16] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study [J]. The Lancet, 2020, DOI: 10.1016/S0140-6736(20)30260-9

[17] Zhao, S, Chen, H. Modeling the epidemic dynamics and control of COVID-19 outbreak in China. Quant Biol, 2020, (8), 11–19.

[18] Tang B, Wang X, Li Q, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions [J]. Journal of Clinical Medicine, 2020, 9(2).

[19] Anastassopoulou C, Russo L, Tsakris A, et al. Data-based analysis, modelling and
forecasting of the COVID-19 outbreak[J]. Plos One,2020,DOI: 10.1371/journal.pone.0230405.

[20]Yang C, Wang J. A mathematical model for the novel coronavirus epidemic in Wuhan, China[J]. Mathematical Biosciences and Engineering,2020,17(3): 2708-2724.

[21]Chen S, Yang JT, Yang WZ, Wang C, Bärnighausen T. COVID-19 control in China during mass population movements at New Year[J]. The Lancet,2020,DOI: 10.1016/S0140-6736(20)30421-9.

[22]Huang R, Liu M, Ding YM. Spatial-temporal distribution of COVID-19 in China and its prediction:A data-driven modeling analysis[J]. Journal of infection in developing countries,2020,14(3):246-253.

[23]Kangkang Wan, Jing Chen, Changming Lu, et al. When will the battle against novel coronavirus end in Wuhan: A SEIR modeling analysis[J]. Journal of global health,2020,10(2).

[24]Geng H, Xu AD, Wang XY, et al. Analysis of the role of current prevention and control measures in the epidemic of new coronavirus based on SEIR model [J]. Journal of Jinan University( Natural Science & Medicine Edition),2020,41(2): 1-7.

[25]Fan RG, Wang YB, Luo M, et.al. SEIR-based Novel Pneumonia Transmission Model and Inflection Point Prediction Analysis [J]. Journal of University of Electronic Science and Technology of China. http://kns.cnki.net/kcms/detail/51.1207.T.20200221.1041.002.html.

[26]Karemera D, Oguledo V I, Davis B. A gravity model analysis of international migration to North America[J]. Applied Economics, 2000, 32(13):1745-1755.

[27]Piore M J. The Dual Labour Market, Theory and Application[M]//Barringer R, Beer S H(Eds). The State and the Poor. Cambridge: Cambridge, Mass Winthrop, 1970:55-59.

[28]Stark O, Taylor JE. Migration incentives, migration types: The role of relative deprivation[J].
[29] Cai F, Wang D. Migration as marketization: What can we learn from China's 2000 census data? China Review, 2003, 3(2): 73-93.

[30] Shen, J. Modelling regional migration in China: estimation and decomposition [J]. Environment and Planning A, 1999, 31(7): 1223-1238.

[31] Henry S, Boyle P, Lambin EF. Modelling inter-provincial migration in Burkina Faso, West Africa: the role of socio-demographic and environmental factors [J]. Applied Geography, 2003, 23(2-3): 0-136.

[32] Fan C. Interprovincial migration, population redistribution, and regional development in China: 1990 and 2000 census comparisons. Professional Geographer, 2005, 57(2): 295-311.

[33] Zhu Y, Chen W. The settlement intention of China's floating population in the cities: Recent changes and multifaceted individual-level determinants. Population Space & Place, 2010, 16(4): 253-267.

[34] Borrow L. School Choice through Relocation: Evidence from the Washington D.C Area [J]. Journal of Public Economics, 2002, 86(1): 155-189.

[35] Statistics Division of Population and Employment, National Bureau of Statistics. Data of sample survey of 1% population in 2015 [M]. Beijing: China Statistics Press, 2016.

[36] National Bureau of Statistics, PRC. China Statistical Yearbook [M]. Beijing: China Statistical Press, 2015.

[37] Investigation Communique of the Ministry of Land and Resources [R]. 2015.