Why the low-skilled workers are reluctant to reallocate:

Evidence From China during COVID-19[*]

Fanghua Li[†]
University of New South Wales

Chenyang Ji
Renmin University of China

Abstract

This paper utilizes the sudden exogenous labor market demand shocks introduced by city-level heterogeneous lockdown and work resumption measures during COVID-19 pandemics, to estimate low-skilled rural workers' reallocation responses to the economic shocks in the short-run, and further explore the mechanism behind the observed behavioral pattern. We utilize the administrative micro level data of the poor population from one county in China which opened open in February 2020 (hence the labor is free to migrate out). This data follows the poor population annually from 2014 to right before the pandemic, and provides extra weekly and monthly monitoring data in 2020. We find the for low-skilled workers whose original migration destinations experience large demand shock, those people are rather inertia with high possibility of staying unemployed in the short run (elasticity around 0.44) and low elasticity of reallocation (around 0.15). Furthermore, females show significantly larger degree of immobility compared to their male counterparts, with smaller chance of getting employed outside of the original destinations and even smaller chance to reallocate to a new city. We continue with the exploration into the mechanism and establish measurements of network dependence based on the concentrate of village-level pre-COVID-19 migration destinations. We find that the high dependence on informal social network largely explains this reluctance to reallocate.

1 Introduction

Whether labor moves perfectly and costlessly across space and why is a crucial assumption in models of spatial equilibrium (Rosen (1979), Roback (1982)).

In terms of the migration cost, extensions based on the standard spatial equilibrium model (Roback (1982)) consider preference shocks for locations (as in Moretti (2012)),

[*]We are grateful for for comments from Moshe Buchinsky. Financial support from the UNSW academic startup funding is gratefully acknowledged.
[†]Corresponding author. Contact email is fanghua.li@unsw.edu.au

transportation cost (as in Morten and Oliveira (2016)), loss in insurance provided by social network (as in Munshi and Rosenzweig (2016)). Determining and quantifying the source of the cost, has clear policy implications and is a crucial step towards welfare improvement. Empirical papers, however, have found that migration responds slowly to economic shocks and that migration rates have been falling in the United States (Molloy et al. (2011)), which is especially concerned for low-skilled and low-income workers since Wozniak (2010) shows high-educated workers are more responsive to the distance labor market boom hence incur less loss.

This paper utilized the sudden and heterogeneous negative labor market demand shocks introduced by lockdown measures during COVID-19 pandemics, to estimate low-skilled rural workers' reallocation responses in the short-run, and further explore the mechanism that block labor's movement across space. We find the for low-skilled workers whose original migration destinations experience large demand shock, those people are rather inertia with high possibility of staying unemployed in the short run. Furthermore, females show significantly larger degree of immobility compared to their male counterparts, with smaller chance of getting employed outside of the original destinations and even smaller chance to reallocate to a new city. We continue with the exploration into the mechanism and find that the high dependence on informal social network largely explains this reluctance to reallocate. Our study provides new evidence for estimations of reallocation responses and add additional micro foundations to spatial equilibrium models (e.g. Coen-Pirani (2010); Kennan and Walker (2011);Allen and Arkolakis (2014) Diamond (2016)).

A key challenge to estimate the responsiveness to local labor demand shock costs has been either the lack of data on detailed migration between locations or the endogeneity of labor demand shock and migration flow. Our paper utilizes the city lockdown and resumption policies in China as an exogenous quasi-experiment and takes advantage of the fact that COVID-19 outbreak right after people returned temporarily for the Spring Festival. In this setup, the labor originated from a early open-up region would experience exogenous and various labor demand shocks from the previous migration destination that they would have been return, and also free to reallocate. We also employ the rich administrative data which follows up people's employment status and migration location during COVID-19 to determine the destination of migration, and the longitudinal data before 2020 to get information of the previous working location and other individual and family characteristics. The quasi-experiment setting and these rich micro data allow us not only to estimate the response elasticity but also discuss mechanisms behind the estimation. In order to encourage migration, those people's migration cost is covered by transportation subsidy. Besides, we also show in the paper that due to the exogeneity of the outbreak of COVID-19, the labor demand shock is also orthogonal to the location-specific utility from amenities and "matching". Hence we argue and prove in this paper that the observed inertia comes from the dependence on the social network and the uncertainty comes from reallocation.

Our paper contributes to several strands of literature.

First, we identify and estimate for the low-skilled workers the short-term responsiveness to the local demand shock. Due to data constraints, most of the spatial, urban and trade literature have made the assumption that the bilateral component of migration costs is zero (e.g. Diamond (2016)) or that labor is immobile across space (e.g.Donaldson and Hornbeck (2016)). A large predominantly empirical literature has focused on the responsiveness of migration to economic shocks. In terms of the sources of economic shocks, Cadena and Kovak (2016) utilizes the geographic variation during the Great Recession and instrument for local labor demand using the standard Bartik (1991) measure that relies on the pre-Recession industrial composition of local employment. Tombe and Zhu (2019) utilizes the cost associated with the hukou registration system and individual data from Chinese censuses to estimate interregional costs of migration. Dix-Carneiro and Kovak (2015) explores the varation caused by regional trade liberalization in Brazil to examin population responses. Bryan et al. (2014) use randomized controlled experiments and find a large migration response when migration is directly subsidized. We follow this line of literature by considering the responsiveness of reallocation induced by the sudden outbreak of COVID-19 and city-level lockdown and resumption policies.

Second, our paper contribute to the discussion of the welfare loss induced by the rural people's dependence on the informal social network. Munshi and Rosenzweig (2016) establishes the model foundation and argues that the rural network stops males of more disadvantaged families from out-migration. Morten (2019) expands the model for understanding the joint determination of migration and risk sharing. In this paper, we instead study the reallocation among the temporary migrants and directly test the relationship between social ties and the responsiveness. We show the social network is also strong among the people already migrate out, especially for females.

Third, since those lockdown policies can be seen as heterogeneous negative labor market shocks and immigration restrictions, our paper also contribute to the literature of the immigrants' location choices. Due to the influx of immigrants to Europe, there is increasing discussion in this topic and most literature in this line discusses the initial location choices instead of the reallocation. In terms of China, Su et al. (2019) argues in China, both job opportunities and amenities play consistent and salient roles in the geographical choice of internal migrants. We show that with high dependence of social network, the immigration shows high level of concentration and high rate of returning to the same locations.

Finally, we also contribute to the increasing studies about the impact of COVID-19, especially on the poor. Whether and how responsive the most disadvantaged people react to this shock not only has strong policy implications but also provides a chance to understand these people better. Current work including Angelov et al. (2021) who utilize the population-wide tax register data to document the impact of the COVID-19 pandemic on firm sales, tax revenues, and sick pay in Sweden and Kugler et al. (2021) uses high-frequency phone surveys conducted by the World Bank and National Statistics Offices in 40 countries and finds larger shares of female, young, less educated, and urban workers stopped working during the pandemic. The advantage of our study is that we our administrative data of the poor population starts in 2014, and follows them up to the most recent, hence compared to survey data, we have little sample selection problem. Besides, our migration data is also administratively collected due to the policies during the COVID-19. This unique data set and also the timing of the COVID-19 in China, which is just around the Spring Festival, gives us a unique chance to study the poor's responses during the pandamic.

The rest of the paper is organized as follows. We first introduce the institutional background during the COVID-19 to show the validity of our quasi-experiment setup and argue the exogeneity of the labor demand shocks for rural migrations. Then in section 3, we give a simple conceptual framework to lead the following discussion. In Section 4, we introduce the data, our measurement of labor demand shocks and in section 5, we formalize our regression setups. We then show and discuss the regression results in Section 6 and explore the mechanism in Section 7. Section 8 concludes.

2 Background and research search

2.1 Locked-down cities and free movement labor

In December 2019, an unknown disease, later named COVID-19, was detected in Wuhan, China (Zhu et.al, 2019 and Lu et. al, 2020). On January 23, 2020, Chinese authorities introduced unprecedented measures to contain the virus, stopping movement in and out of Wuhan, and 15 other cities in Hubei province — home to more than 60 million people. Other cities in China have also adopted dramatic measures to reduce human interaction, including enforcing strict quarantines, prohibiting large-scale private and public gatherings, restricting private and public transportation, encouraging social distancing, imposing a curfew, setting up checkpoints around and even locking down entire cities. Fang et.al (2020) summarizes the various forms of population mobility control in different cities, and states that 115 cities among all 293 cities issued different levels of lockdown policies until February 29, 2020. In Figure 1, we show the spatial distribution of various levels of city lockdown before February 10[1]. As we can see, besides Hubei Province, where Wuhan locates, almost all southeastern cities which are the destinations of the predominant migration flows in China (from the inland to the southeastern coast)[2], were locked down to certain levels in February 2020.

On February 10, after the virus been put into control for most of China, almost all of provinces (except Hubei) announced work resumption for most of the industries except for restaurants, cinemas and other entertainment industries[3] and gradually lowered down the response level to epidemic situation. Appendix Figure 14 shows the timing for each province to lower down their response level. As we can see, there is a great variation in

[1]We choose February 10 as cutoff since most cities start work and production resumption at that day

[2]See Chan (2012)

[3]The new confirmed cases outside of Hubei Province on Feb 9 dropped under 500. Data sources: https://www.sohu.com/a/371996711_428505

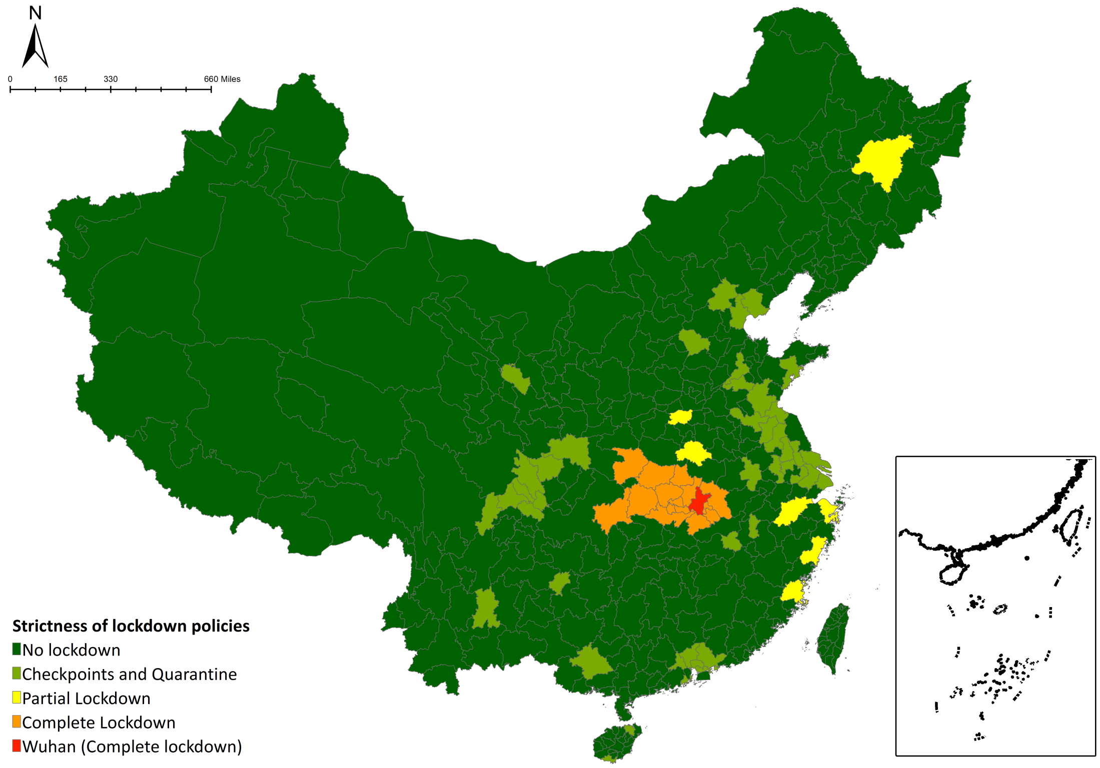


Figure 1: Cities with Control Measures in China, February 2020

terms of the timing and extension of city reopening. Beijing lowered down from the first response level to the second as late as April 30, 2020 while Guangdong, another major host of rural migrants, was among the first to lower down the response level on February 24. Most of the migration origin provinces including Henan, Sichuan and Anhui, lowered down at the end of February, suggesting little restrictions of labor's out-migration. On the other hands, each local governments adopted various policies to help firms to go through the financial difficulties during the pandemic, including tax exemption, lowering the loan interests for small firms and helping staffs to return through special trains or airlines for point-to-point transportation. Zhang et.al (2020) summarizes the encouragement policies and points out that there is a also large variation among cities. Due to the government-imposed policies, the major destinations of the rural migrants in China experienced various labor demand shock determined by the relative relationship between the strictness of the lockdown and the generosity of production resumption stimulation policies.

On the contrary to the locked-down cities, the low-skilled labor from certain parts of rural China are free to move. This is due to the fact that the outbreak of COVID-19 coincided with the 2020 Chinese Spring Festival (the announcement of the discovery of a new unknown virus was 5 days before the festival[4] and the Wuhan lockdown was 2

[4]The official announcement was made by a renowned Chinese epidemiologist Dr.Nanshan Zhong, who discovered the SARS coronavirus in 2003, source: http://www.xinhuanet.com/2020-01/20/c_1125487200.htm in Chinese

days before the festival), and almost all rural migrants already left the city before the escalation of epidemic policies. China's Spring Festival is the most celebratory time of the year in China, during which individuals travel back to their hometowns, or in more economics terms, that the migrants temporarily leave the destination cities and go back to their origins. According to the official records, during the 15 days pre-holiday in 2020 (Jan 10-24, 2020), people in China made 1.139 billion trips, 2% higher than that in 2019 and were estimated to make over 3 billion trips during the 40-day pre and post holiday, or Chunyun[5].

We take Beijing as an instance, the capital and a megacity hosting 3 million rural migrants in China. This amount of migrants accounted for 13% of Beijing's permanent residents and formed the main low-skilled labor force for Beijing's economy[6]. Figure 2 depicts the migration flow out of Beijing 25 days before and after the Spring Festival in 2020 (January 1,2020 to February 10,2020), compared with that of same period in 2019. We can see that in 2019, the scale of out-migration kept increasing till the Spring Festival Eve, which captures the home-return flow of the migrants, and the net-inflow after the Spring Festival is almost the same scale of the net-outflow before. In 2020, before the official announcement of the discovery of a new unknown virus, the migration trend in and out of Beijing is comparable in the two consecutive years even in slightly higher amount[7]. This suggests that the outbreak of COVID-19 didn't stop the return migration due to the coincident timing. After the 15-day holiday, those migrants would've been once again return to cities for work, but interrupted by the sudden outbreak of the virus and nation-wide city lockdown policies. By Feb 14, 2020 (20 days after Spring Festival), the size of the return migrants of the whole country is only 80 millions, far below that in 2019[8]. Beijing is almost the cities with the strictest and longest city lockdown, hence we can see the migration flows flatten out to almost zero after the festival.

Our study takes the advantage of the fact that while the major migration destination cities experienced various degree of labor demand shock, the low-skilled rural labor were free to make migration and reallocation choices in certain areas of China (the migration origins), all due to the coincidence timing of the COVID-19 outbreak and the Chinese Spring Festival. Hence we can utilize this exogenous and heterogeneous labor demand shocks as a quasi-experiment, to examine how responsive the free movement labor can adjust to the demand shock.

2.2 Low-skilled workers from rural China

In this subsection, we mainly argue that this demand shock is mainly for the low-skilled

[5]Data from Ministry of Transport of the People's Republic of China: Big data! The travel volume predictions during Lunar New Year holiday in 2020 http://www.mot.gov.cn/fenxigongbao/yunlifenxi/202001/t20200109_3322161.html (Jan 9, 2020), Accessed 15th Feb 2020 ([in Chinese]).

[6]Data source from Beijing government, http://news.yktworld.com/2019/20200609085340164l641.html

[7]The relative higher amount of out-migration around 20 days before the Chinese Spring Festival in 2020 is the 2020 New Year, hence relatively higher than that in 2019

[8]Data source from the press conference of the State Council

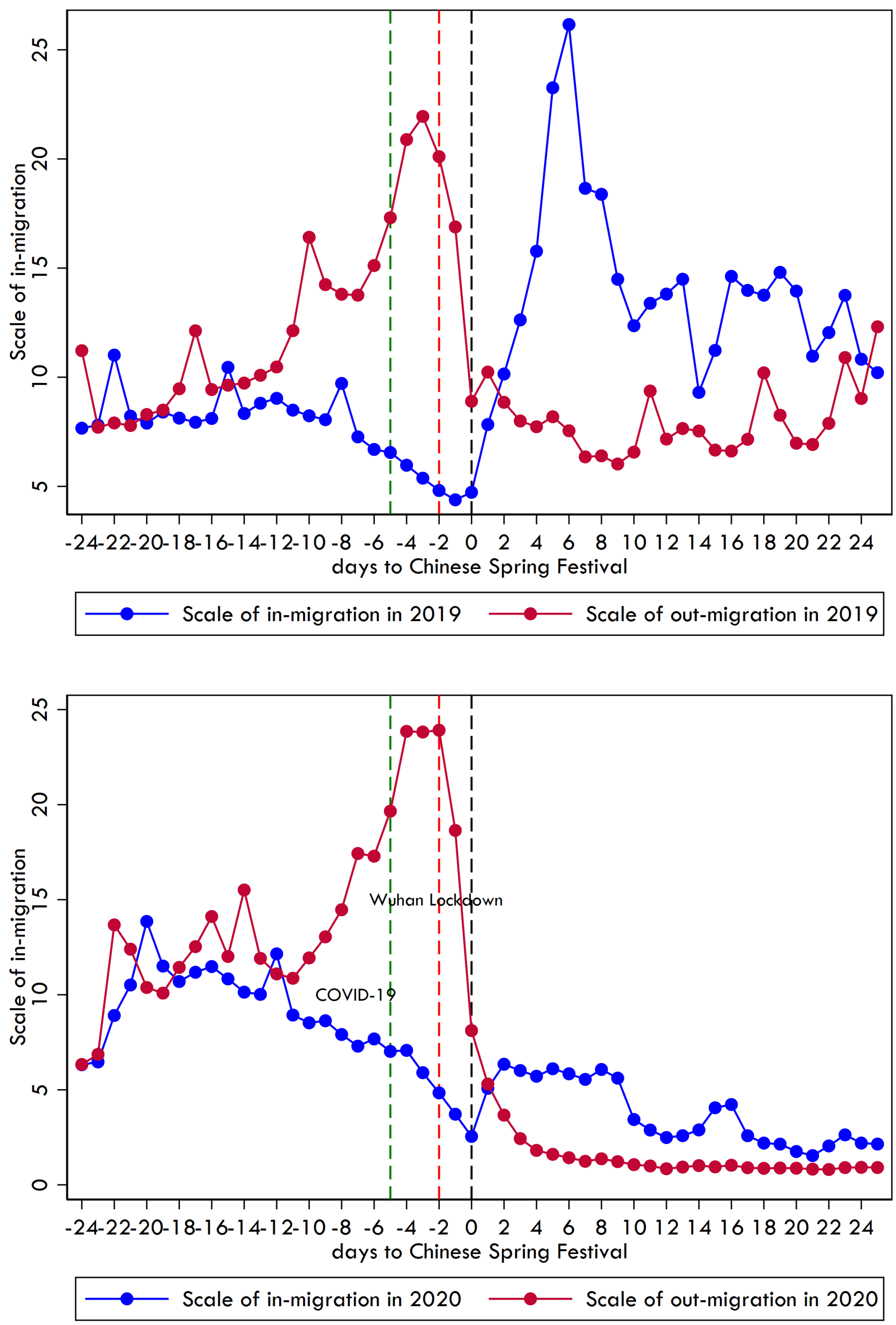


Figure 2: Migration flow from Beijing, 25 days before and after Chinese Spring Festival in 2019 and 2020

workers migrated from rural China and our research design also works the best for this specific group. In 2019, the overall size of rural migrant workers is 290 million[9] and takes up to one quarter of the whole population in China.

The outbreak of COVID-19 witnesses a shift to work from home due to social distancing requirement around the world. Various studies including Mongey et al. (2020), Gallacher and Hossain (2020) study the feasibility of working remotely and suggest that poorer workers, workers with lower education and part-time workers tend to be employed in jobs for which remote work is less possible in developed countries like US and Canada. They also suggest that workers in occupations for which the possibility of remote work is less likely experienced larger employment losses.

In other words, only for people employed in occupations which are infeasible for remote working, the city lockdown policies can be interrupted as the labor demand shock. According to Bureau of Statistics of China, the rural migrants are mainly occupied in labor-intensive industries, with 18.7% work in constructions, 12.3% in services and 12.0% in transportation and delivery industries. Those people are also low-educated, with 15.3% elementary school graduates, 56% middle school graduates, 16.6% high school graduates and only 11% of them with three-year college diploma and above. We conducted a survey on March 26, 2020 in Xin County, Henan Province, which is one inland county of China, for all the rural migrants that used to work outside of the county in 2019. The Figure **??** depicts the industry distribution of those rural workers. As we can see, the industry distribution is very similar to the overall statistics in China, with most of those people are occupied in industries that require on-spot operation, with the construction and interior decoration industries take over 29%. Hence for this specific group of people, the city lockdown can be equalized to labor demand shocks.

On the other hand, those low-skilled rural migrants are free to reallocate for two reasons. First, those low-skilled rural workers often work under no contract. According to a national survey, only 35.1% rural workers in China work under contract[10]. Take rural workers in the construction industry for instance, according to Pan et.al(2014), the national average contract rate is only 17.4%. Even for people with contract, the contract often has no fixed term or less in a year (12.4% have contracts with no fixed term and 4.2% with contract term less in a year). According to our survey, only 8.6% rural migrants have the full coverage of insurance by the employed firms, suggesting a low long-contract rate among these people. Second, the rural workers in our sample are entitled to transportation subsidies up to 1,000 CNY which is enough to cover the transportation costs to almost anywhere in China (the highspeed train from this region to Beijing is only around 400 CNY).

To sum up, the returned rural migrants before the city lockdown can be seen as peo-

[9]Data from Bureau of Statistics of China: http://www.stats.gov.cn/tjsj/zxfb/201704/t20170428_1489334.html

[10]Data source: https://www.yicai.com/news/5303859.html in Chinese

ple facing various exogenous labor demand shock and also free to make adjustments accordingly. Hence we'll look into their choices and ask the question how responsive they can be to the demand shocks.

3 Conceptual framework

We consider here a discrete choice model of location and employment decisions for rural workers, which is a simplified version of the model considered in Buchinsky et.al (2014). In each period, rural worker chooses employment status (whether stay in the agricultural sector or not), the location of employment (local within commuting distance or migrate out to one of the other cities in China in order to maximize the expected discounted present value of remaining lifetime utility. In other words, the mutually exclusive choice set of employment status options has size of $I+2$, including non-employment and staying in agricultural sector ($k=0$), temporary migration for employment ($k=1,\ldots,I$) and local employment ($k=I+1$). The first and the last options correspond to location $r=0$ and the migration option to location $r=i\in\{1,\ldots,I\}$, with each i represent one city in China. In order to control for heterogeneity in per-period utility and job opportunities and other unobserved factors, we assume that there are J fixed discrete types of individuals (males vs. females, or educated vs. low education). The per-period utility for each option k of individual i, of type j, at time t, in region r ($r=0,1,\ldots,I$), is given by:

$$u^j_{ikt}=\begin{cases} b^j_{kt}(x_{it},\epsilon_{it})+\tau^j_r(x_{it},\mu_{ir}), & k=0\\ w^j_{rt}(x_{it},\epsilon_{it})e^{iR't}+\tau^j_r(x_{it},\mu_{ir})-\gamma_j\mathbf{I}(r_t\neq 0), & k=1,2,\ldots,I\\ w^j_{rt}(x_{it},\epsilon_{it})e^{it}+\tau^j_r(x_{it},\mu_{ir}), & k=I+1\end{cases} \tag{1}$$

in which b_{kt} represents the pecuniary benefits for $k=0$ in period t. If $k=0$ (staying in the agricultural sector), then b^j_{0t} is the per-period home production value and leisure value from staying in the agricultural sector, while if $k>0$, then w^j_{rt} is the deterministic components of the wage offer function in region r, corresponding to the individual i's characteristics x_{it} and stochastic term $\epsilon_{iR't}$, with R' as the past working location till $t-1$. In other words, we allow the possibility that by reallocation, the pecuniary gain from work would incur larger uncertainty (couldn't find the best suitable job, larger expenses etc.). We assume each type j has different wage offer function. For staying in the local, we assume the stochastic part is smaller and not path dependence. τ_r represents the non-pecuniary benefits or per-period preference for residing in region r, varies by the location r and individual type j. This preference is a function of the individual's characteristics, x_{it}, as well as an idiosyncratic term μ_{ir}, which represents the "match qualities", and captures the migrant's valuation of regional amenities and/or policies. γ_j is the location and type specific costs incurred when switching from migration.

Although the model does not impose any restrictions on the choice of employment location and option, there are natural restrictions placed on the choice of employment sector and region of employment. That is, a job in each region is in the individual's choice set only if an employment offer is received. We assume that, in each period t, the probability of receiving an offer in location r, at time t, for type j is P^j_{rt}.

Since we only care about the short-term transition behavior, hence a partial equilibrium model is enough in which housing prices and wages are taken to be exogenous. To solve the large dimension problem, conditional on the optimal choice ${k^*_{i,t-1}}^j \in [1, \ldots, I]$ in period $t-1$ the employment status options $k^*_{it,j}$ can be collapsed into a vector S of size 4:

$$s^j_{it} = \begin{cases} 1, & k^{j*}_{it,j} = k^{j*}_{i,t-1} \in [1, \ldots, I] \quad \text{unchange} \\ 2, & k^{j*}_{it,j} = 0 \quad \text{switch to unemployment} \\ 3, & k^{j*}_{it,j} = I+1 \quad \text{switch to local employment} \\ 4, & k^{j*}_{it,j} \in [1, \ldots, I] \neq k^{j*}_{i,t-1} \quad \text{switch to another city} \end{cases} \tag{2}$$

including unchanged employment status $k^{j*}_{it,j} = k^{j*}_{i,t-1}$, switch to unemployment $k^{j*}_{it,j} = 0$, to local employment $k^{j*}_{it,j} = I+1$ and to another city $k^{j*}_{it,j} \neq k^{j*}_{i,t-1} \& k^{j*}_{it,j} \in [1, \ldots, I]$. In each period t, individual i makes optimal choice $s*$, given the exogenous offer arriving rate vector $[P_0, P_1, \ldots, P_I]^j_t$, wage offer vector $[w_0, w_1, \ldots, w_I]^j_t$, the non-pecuniary function vector $[\tau_0, \tau_1, \ldots, \tau_I]^j_t$ and individual i's characteristics x_{it}, and the stochastic terms $\epsilon_{iR't}$ and μ_{ir}. We take it as granted that the local wage w_0 is lower than the migration wage w_r for all type j.

If we have a vector of random city-level labor demand shocks, paired with the labor's employment and location choices in two consecutive periods of the whole country, we can estimate the new equilibrium in the second period. However, due to the fact that, we only have the labor's choices from one typical pro-migration county, hence based on this framework, we're interested in estimate the elasticity of type j's employment response s^* to the change of job offer arrival rate (or working opportunity) in location $r^{j*}_{i,t-1}$, i.e. $P_{r^*_{t-1}}$. In particular, we're interested in the elasticity of switching,

$$\varepsilon^j_{s=3|4, P_{r^*_{t-1}}}|_{w,\tau} = \frac{\Delta Pr(k^{j*}_{it,j} \in [1, \ldots, I+1] \neq k^{j*}_{i,t-1}) / Pr(k^{j*}_{it,j} \in [1, \ldots, I+1] \neq k^{j*}_{i,t-1})}{\Delta P_{r^*_{t-1}} / P_{r^*_{t-1}}} \tag{3}$$

Another side of the same coin is to estimate the elasticity of unemployment due to the labor market shock, i.e.:

$$\varepsilon^j_{s=2, P_{r^*_{t-1}}}|_{w,\tau} = \frac{\Delta Pr(k^{j*}_{it,j} = 0) / Pr(k^{j*}_{it,j} = 0)}{\Delta P_{r^*_{t-1}} / P_{r^*_{t-1}}} \tag{4}$$

When we have exogenous shock $\Delta P_{r^*_{t-1}} / P_{r^*_{t-1}}$, we can estimate the labor's reallocation elasticity in the short-term. To be specific, if our labor demand shock LMS, is also exogenous with individuals' location preferences τ^j and the reallocation fixed cost γ^j is covered, we can estimate the elasticity of reallocation just due to the uncertainty induced by $e^{iR't}$.

If the labor reallocation is inelastic, i.e., $\varepsilon_{s=3|4} \to 0$ and $\varepsilon_{s=2} \to 1$, it would suggest large expected loss due to reallocation for this group of people. While the larger the elasticity of reallocation, the more flexible the labor's reallocation, and the smaller the market friction should be.

4 Data and measurement of demand shocks

4.1 Data source and descriptive statistics

Our data combines the administrative data and survey data from Xin County, one county from inland Henan Province, which can be seen as a representative county in China. With a residential population of 26 million and registered population of 38 million, Appendix Figure 15 shows that, in terms of economics development level (per capital GDP in 2016), urbanization rate (ratio of urban Hukou status or ratio of employed people in agricultural industry), population density (per capita per square meter) and education level (average years of education), Xin County is around the median level of all the above statistics among the 2,000 county-level administrations[11]. On the other hand, in terms of migration, Henan is famous for its rural labor exporting, and Xin County is no exception. From Figure 16(b) we can see that, Xin County, as a typical county in Henan province, witnesses a much larger proportion of children under age 14 and senior population above age 65 compared to the county average or county median. The ratio of people who choose to out-migrate to other places and work temporarily account for 22% of the total registered population, which is below the lowest 10% in all county-level administration.

During the COVID-19, Xin county is classified into low-risk regions as early as March 8, and work resumption is among the top priority of the provincial government's job[12]. Hence, Xin County is a representative county in terms of economic development, education level, urbanization rate and population density, while a typical county in Henan in terms of out-migrating and the rural migrants were free to move starting from early March in 2020.

We built up our micro dataset by linking data from the following two sources:

- **Administrative panel data for the poorest 15% rural families**
  Starting from 2014, China started the Targeted Poverty Alleviation (TPA afterwards) program aimed at ending poverty by the end of 2020. The Leading Group Office of Poverty Alleviation and Development (of the State Council) (LGOPAD) built up a nationwide registration system collecting individual level data (containing 29 million households and approximately 90 million individuals) and updating annually (quarterly for several years) starting from 2014 till now.

[11] GDP and size of administration data come from the latest CHINA STATISTICAL YEARBOOK(TOWNSHIP), 2016. All the residential population data comes from the latest census, which took place in 2010.

[12] Data source: https://www.sohu.com/a/378708162_100133062

Xin County hosts 12 thousand poor families and 45 thousand poor people under this TPA program, which can been seen as the poorest 15% rural people in this County[13]. We are authorized to use these longitudinal micro data for the whole county, containing information like assets, households income, family composition and labor supply etc.

In terms of the labor supply data, before 2019, the information including the working location, months in work and the salary income is self-reported with proof document provided by firms. For the 2019 and 2020 information, due to the COVID-19 prevention policies, all cross-city travel data is recorded by authority, hence the working location and the starting month to work is also administrative recorded data. Furthermore, starting from 2020 March, the local government also collectted the weekly employment information till April for six consective weeks. Hence we get the monthly employment information for the whole year and also weekly employment information for March and April.

- **Survey Data**
  The administrative data has certain advantages, for instance, it's more accurate since it's government verified, but on the other hand, it's limited in terms of variables, for instance, social network, attitudes and expectations, contract details etc. Hence we conducted a Labor Force Survey in Feb 2020, via the help of the local government and officials and mainly focus on the insurance, industry, skills, channels to get work etc.. The detailed of the survey questions are provided in Appendix II.

  We link this survey data with our administrative data set via uniquely identified national ID.

Appendix Table 7 summarizes the demographic statistics for the rural people in our data set. In 2019, among the people in the working ages (16-65 years old), 42.72% females and 70.99% males participate in paid job. The participation rate by gender and age are shown in Appendix Figure 17. Besides, there are high proportion of them migrate to big cities to work. As shown in Appendix Figure 18, 72.14% of working females and 80.97 of working males work outside of the county. In terms of migration destinations, as shown in Figure 3, the bigger the city, the more attractive to rural migrants. Besides, the distance from home is also another determinant, but this factor is only significant among the smaller cities. Furthermore, the rural workers from Xin county work in 242 cities among all 292 cities of China, which makes it possible for us to explore the variation in terms of demand shock in different cities and is also another piece of evidence supporting the free movement of those people. In terms of the top destinations, Beijing and the surrounding big cities take up 12.46%, Yangtze River Delta City Group (around Shanghai) takes up 13.52%, Pearl River Delta (around Shenzhen) takes up another 13.76% and Wuhan takes on around 5%.

[13] For the details of TPA and the identification problem, see Li et.al(2021)

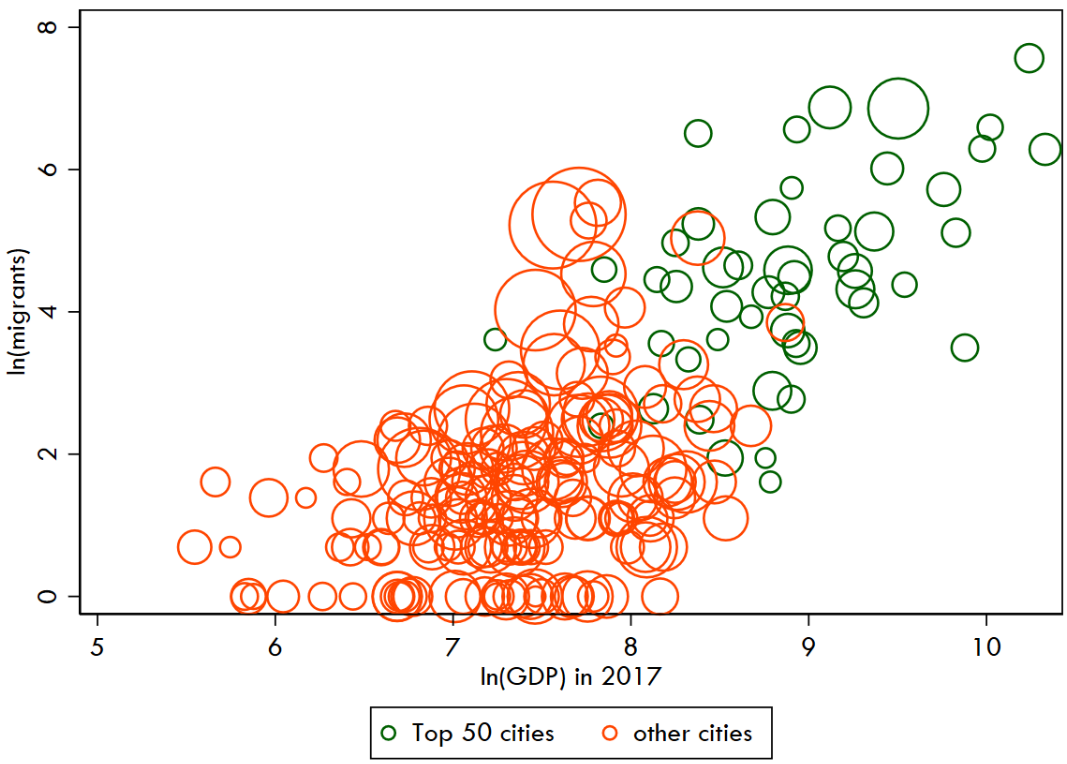


Figure 3: Migration location distribution of rural migrants in 2019, from Xin County

4.2 Measurement of labor demand shock for low-skilled labor

In this section, we introduce our measurement of the labor demand shock for low-skilled labor, and justify the validity of it. We utilize the daily migration data from the Baidu Qianxi web platform, which provides daily travel flow index among cities and the traffic intensity within each city, based on the mobile geolocation data from Baidu Inc., the largest search engine firm in China.[14]. This migration data has been used by Fan et.al (2020) to quantify the causal impact of the Wuhan lockdown, on the spread of COVID-19 in China.

As we state in the background section, the rural workers should've been return to destination cities after the Spring festival, but interrupted by the lockdown policies. Inspired by this direction of migration after the Spring Festival, we can construct a measurement of city low-skilled labor demand at Day T based on the size of the return labor, i.e., the accumulated net inflow after the Spring Festival till T. To be more specific, for city i at Day T after the Spring Festival, we define the labor demand as the following Equ.5:

$$D_{i,T} = \sum_{t=0}^{T} \Delta \text{inflow}_{i,t} = \sum_{t=0}^{T} [\text{inflow}_{i,t} - \text{outflow}_{i,t}] \tag{5}$$

We can then then define the labor demand shock, as the actual labor demand in 2020 at day T relative to the expected labor demand at the same time:

$$\Delta D_{i,t} = \frac{D_{i,T} - \bar{D}_{i,T}}{\bar{D}_{i,T}} \tag{6}$$

with $\bar{D}_{i,T}$ defined as the expected labor demand for city i on Day T.

There are several different ways to calculate the expected labor demand at day T (or expected accumulated net inflow after Spring Festival till day T), one by assuming that the net inflow is comparable in the two consecutive years, and the other assuming that the net inflow ratio before and after the Spring Festival is comparable in the two consecutive years. To be specific, $\bar{D}$ can be defined as

$$\bar{D}^1_{i,t} = \sum_{t=0}^{T} \Delta \text{inflow}^{2019}_{i,t} = \sum_{t=0}^{T} [\text{inflow}^{2019}_{i,t} - \text{outflow}^{2019}_{i,t}] \tag{7}$$

or

$$\bar{D}^2_{i,t} = \sum_{t=-24}^{t=0} \Delta \text{outflow}^{2020}_{i,t} \times \frac{\sum_{t=0}^{T} \Delta \text{inflow}^{2019}_{i,t}}{\sum_{t=-24}^{t=0} \Delta \text{outflow}^{2019}_{i,t}} \tag{8}$$

Again, the above measurement based on migration data can only capture the labor demand that requires on-spot work, hence covering most of the low-skilled jobs, like construction,

[14]The website is http://qianxi.baidu.com/. The data comes from location data from users of Baidu-related products, including but not limited to search engine (Baidu search), map (Baidu Map), food delivery (Baidu Delivery), travel (Xiecheng) phone applications, and other firms' applications which invoke the map API from Baidu. Inc. Based on the *China Mobile Internet Research Report in 2013-2014*, Baidu Inc. takes 63.7% of all phone map users.

manufacture, and low-skilled service sectors.

To prove the validity of our measurement, we collected the local government's lockdown policies from news media and government announcements for two major cities in China with completely different resumption strategies - Beijing and Shenzhen[15]. Table 8 summary the policy differences in various aspects[16]. Shenzhen is the among the first in China to open up. On February 24, any people (except from Hubei Province) can travel to Shenzhen, without quarantine, and all public transport went back to normal on the same day. On the contrary, Beijing lowered the prevention level as late as the end of April, while the public transport capacity limit was still in effect (65% of the full capacity) at that time. Higher level of lockdown and prevention policies means it's difficult or even impossible for the return-migrants to go back to cities for work. On the other side, Shenzhen implemented 18 policies aiming at stimulating production resumption, ranging from rent exemption, to better government services to firms. Till February 20, over 60% state-owned firms already went back to normal[17]. On the contrary, Beijing only exempt the rent for February for some firms and others with 50% discount. To show these different lockdown policies indeed expose different labor demand shocks, we show the time trend of our measurement of labor demand shocks for these two cities in Appendix Figure 19. As we can see, our measurement for Beijing stays around -0.6 till the end of March while Shenzhen picked up after the prevention level-down and gradually went back to around 0 shock till then.

4.3 Sample Selection

We construct daily measurements for each Chinese city till the end of April in 2020[18]. There is a potential problem with the previous measurement, that if the city used to be pro-migration city, with negative net inflow, and the lockdown policy is strict and prevents people from out-migration, then the above measurement will be strongly negative but not necessarily capture the actual size of the demand shock. To avoid this false measurement, we adopt two strategy. The first is we exclude those with the magnitudes of shocks larger than 1. In principle, it's not possible to have a -100% negative shock and it's also rare to have a demand boom larger than 100%. The second strategy, is to limit our discussion to the big cities that famous for in-migration. We follow the classification of Chinese city level in 2019, and select only the 4 mega cities, 15 Tier 1 cities and 30 Tier 2 cities.

In Table 1, we report the descriptive statistics and the pairwise correlation coefficients between these two measurements. As we can see, in February, the negative shock is big

[15]Shenzhen is a megacity hosting 8.5 million domestic migrants. This population accounted for 65% of its residents and formed the main labor force for Shenzhen's economy.

[16]Summaried by the author. Data sources: http://www.chinawuliu.com.cn/zcfg/202003/07/495096.shtml, https://baijiahao.baidu.com/s?id=1661951434507619859&wfr=spider&for=pc

[17]Data sources: https://baijiahao.baidu.com/s?id=1659414970379936562&wfr=spider&for=pc

[18]On March 23, 2020, for the first time, there was no new case reported in China. As early as February 10-17, 2020, there were only few new cases reported each day, hence the central government suggested all areas to adjust the lockdown policies accordingly and prepare for work resumption.

(around -64%–71%) but with smaller standard variations among cities, in March, due to the differential lockdown policies, the mean of the shock shrunk to around -35% among the big cities and -40% among the overall cities. Besides, the two measurements are highly correlated with one another but not perfectly collinear, providing us with a nice robustneses check.

	Feb 10 - Feb 28		March 1 - March 31	
	Mean	SD	Mean	SD
Sample 1:Top 49 cities and $\Delta D\% \in (-1,1)$				
$\Delta D_1\%$: expected demand defined by Equ.7	-0.649	0.249	-0.349	0.276
$\Delta D_2\%$: expected demand defined by Equ.8	-0.716	0.235	-0.463	0.290
$corr(\Delta D_1\%, \Delta D_2\%)$	0.984		0.889	
Sample 2: $\Delta D\% \in (-1,1)$				
$\Delta D_1\%$: expected demand defined by Equ.7	-0.641	.276	-.405	.321
$\Delta D_2\%$: expected demand defined by Equ.8	-0.705	0.252	-0.513	0.311
$corr(\Delta D_1\%, \Delta D_2\%)$	0.827	0.849		

Table 1: Descriptive statistics and correlation between the two measurements, for different periods and different samples

Figure 4 depicts the histogram distribution of $\Delta D1\%$ at the end of March, with Beijing, Shenzhen, Shanghai and Hanghzou highlighted. As we can see, Beijing is among the biggest 25% labor demand shock, with the accumulated inflow in the post-holiday period shrunken by around 50%, while for Shenzhen and Shanghai, the demand shocks are among the lowest 10%, with the labor demand increases at around 10% at the end of March.

4.4 Stylized facts

Before we turn to rigid regression setup and estimation, we'll provide several stylized facts to initiate our study.

In Figure **??**, we can see that by the end of March, the people under employment is significant lower than that of 2019. We depict the overall employment rate for working age (16-74) in Xin County up till April 1 2020. We can see that, if the public positions provided by the local governments are excluded, the employment rate drop sharply.

On the other side, for the people who got a job by then, we can see in Figure 6, that those people more tend to work in their hometown instead of migrating out.
With the large labor shortage in the cities with almost no restrictions on movement, the salary those people received in 2020 is getting lower. In Figure 7 and **??**, we show that the

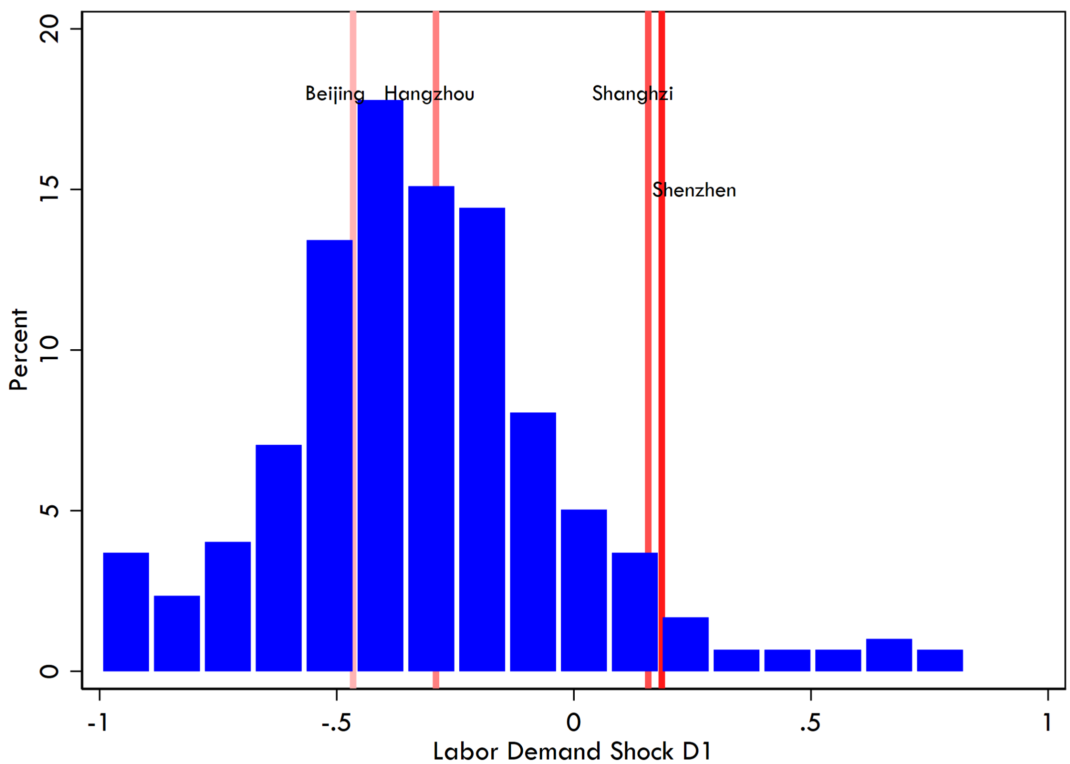


Figure 4: Distribution of $\Delta D1\%$ on March 31, 2020

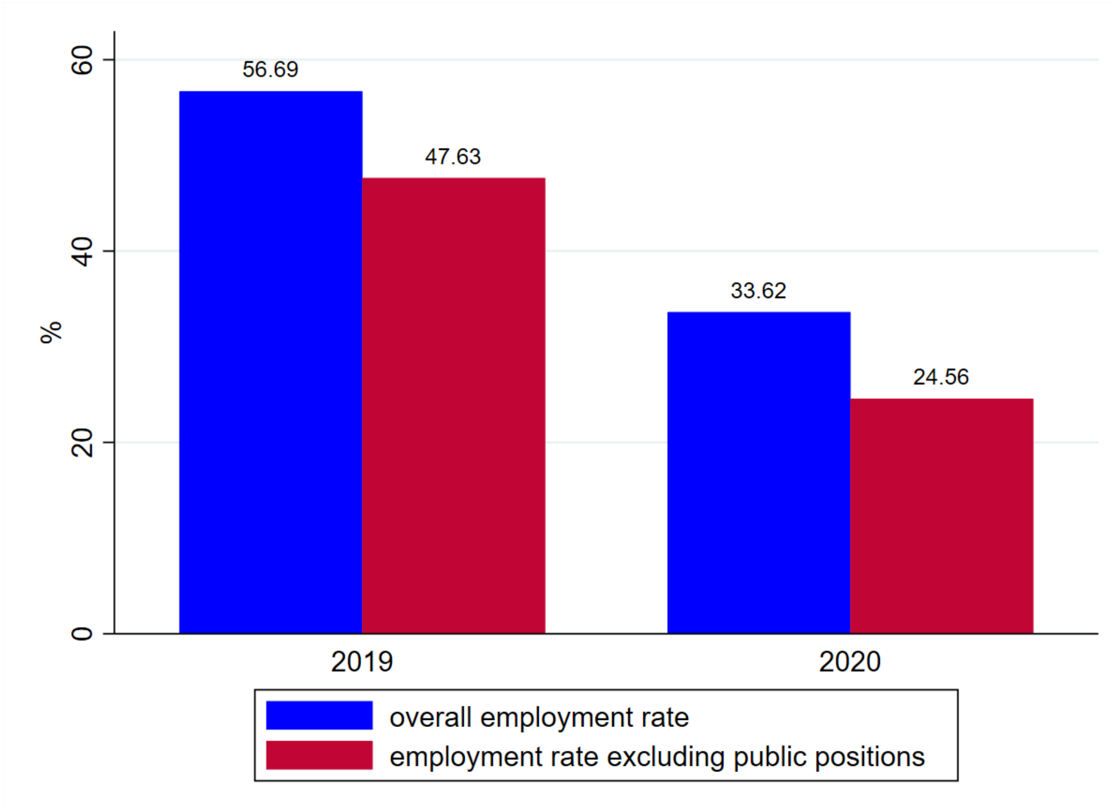


Figure 5: Comparison of the employment rate for working age (16-74) in Xin County

the average wage decrease sharply between 2019 and 2020, one reason maybe due to the fact that more people tend to work in the local hence experiencing lower salary level.

5 Empirical Settings

In this part, we'll specify the regression models in this paper. We'll first introduce our basic setup in which we believe exogeneity assumption is valid, and then continue with the discussion of potential endogeneity and our solution for this.

5.1 Basic setup

To examine the responsiveness of the low-skilled rural labor force to cities' labor demand shock, we estimate a model of the individual choice problem in Equation **??** using conditional logit representation (McFadden 1974). Logit model allows each individual to choose from among J unordered choices $1, 2, \ldots, J$. Unlike Wozniak (2010) or Diamond(2016) which estimate the spatial equilibrium structurally, our paper focuses on the question how the low-skilled workers can adjust smoothly to the local labor demand shock to avoid losses, therefore our choice set is not the city set. We're interested in the two dummy

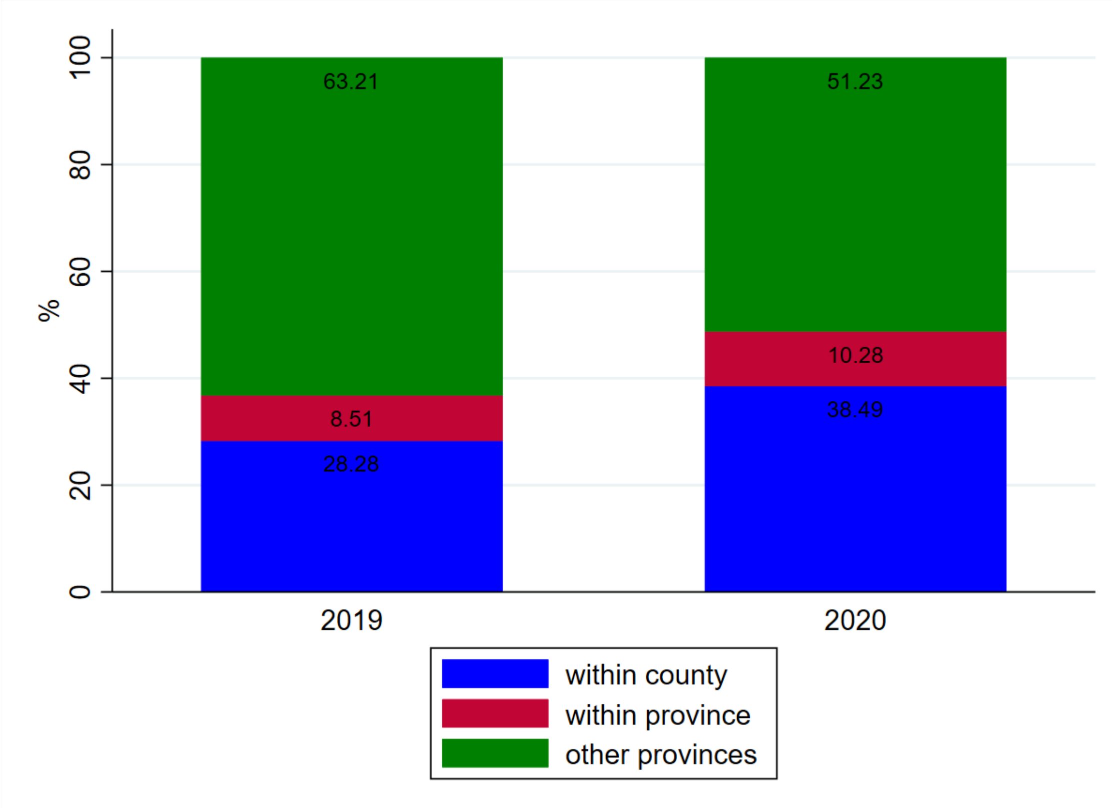


Figure 6: Compare the location type in 2019 and 2020

variables:

$$reallocate_i = \begin{cases} 1 & \text{if } i \text{ choose to reallocate and migrate to other cities} \\ 0 & otherwise \end{cases} \tag{9}$$

and

$$unemploy_i = \begin{cases} 1 & \text{if } i \text{ stay unemployed} \\ 0 & otherwise \end{cases} \tag{10}$$

The conditional probability of observing $y = 1$ with $y \in \{\text{reallocate}, \text{unemploy}\}$ is the following:

$$Pr(y = 1) = \frac{\exp\{\mathbf{X}'\beta\}}{1 + \exp\{\mathbf{X}'\beta\}}$$

Among the covariates $\mathbf{X}$, the main explanatory variable is the measure of previous migration city's labor market shocks (LMS) defined in the previous section. To be specific, based on our description previously, if not for the Spring Festival celebration, the individual i would've been stayed in the destination location r chosen in 2019, hence we assume an individual i experienced the shock from the city he/she used to work in before the 2020 Spring Festival, i.e., $LMS_i = LMS_r$.

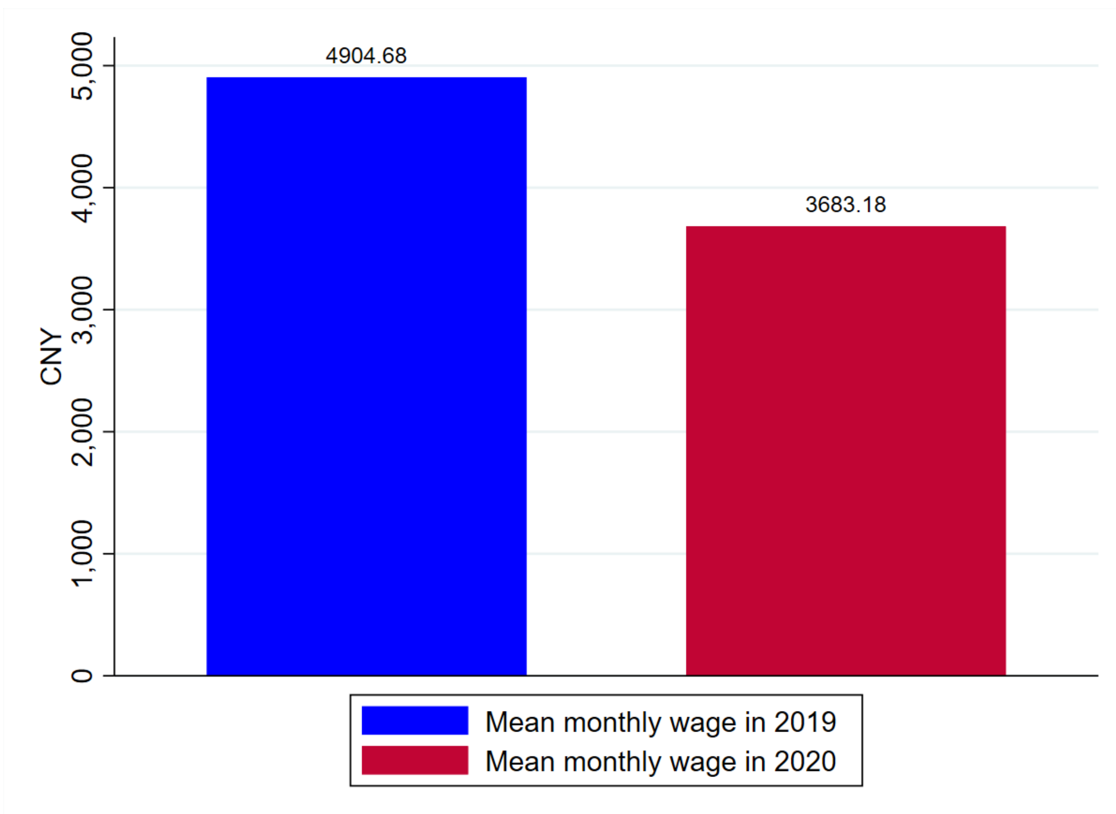


Figure 7: Average wage in 2019 and average expected wage in 2020

We also include the following controls: individual characteristics like gender, education level and age and family demographic information including number of children under 5-year-old, number of students in schools, number of adult females and adult males, number of sick people and old people in the family, also the average per capita income between 2016-2019. To control for the potential correlation, all regression are run with standard errors clustered by village.

The coefficient vector β is estimated by maximizing the following probability of observing y_i using the following log-likelihood function:

$$\log L(\beta) = \sum_{n=1}^{N} \{reallocate_i \times \log(\pi_i) + (n_i - reallocate_i) * \log(1 - \pi_i)\}$$

where $\pi_i = Pr(y_i = 1)$ depends on the covariates X_i and a vector of p parameters β and the data contain observations on the choices of N individuals.

5.2 Extensive model

To understand people's behavior patterns and the mechanisms better, we then extend the previous model to include the full set of options to avoid the labor demand shock. Other than voluntarily or involuntarily exit the labor market in temporary (and bear the pecuniary loss), individuals have two options, one is to reallocate as we stated previously, the other is to switch to work in the local with significantly decreased salary (*local*). Hence, we utilize Multinomial Logit model to estimate the full option set for the employed individuals:

$$employ_i = \begin{cases} 1 & \text{if } i \text{ choose to go back to the original city} \\ 2 & \text{if } i \text{ choose to work in local and not migrate} \\ 3 & \text{if } i \text{ choose to reallocate and migrate to other cities} \end{cases} \tag{11}$$

then the conditional probability of observing $employ_i$ is the following:

$$Pr(employ_i | \sum_{j=1}^{3} employ_{ij} = 1) = \frac{\exp(\sum_{j=1}^{3} employ_{ij} x_{ij} \beta)}{\sum_{di \in S_i} \exp(\sum_{j=1}^{3} d_{ij} x_{ij} \beta)}$$

where S_i is the set of all possible $employ_i$ outcomes, which are denoted d_i. The coefficient vector β is estimated by maximizing the conditional probability of observing $employ_i$ using the following log-likelihood function:

$$\log L(\beta) = \sum_{n=1}^{N} (\sum_{j=1}^{3} employ_{ij} x_{ij} \beta - \log(\sum_{di \in S_i} \exp(\sum_{j=1}^{3} d_{ij} x_{ij} \beta)))$$

where the data contain observations on the choices of N individuals. Similarly to the basic setup, on top of the main explanatory variable $LMS_i = LMS_r$ for individual i used to work in location r, we also include individual and family characteristics in the covariates X.

5.3 Instrumental Variable

In the previous setups, we directly examine the relationship between our defined labor market shock (LMS) and people's responses, under the assumption that the LMS is exogenous. In principle, since people cannot influence how destination cities choose their lockdown policies, it's plausible that the LMS is exogenous. However, if people sort themselves around the destination cities in 2019 according to some unobserved preferences or abilities (included in $\tau_j(x_{it}, \mu_{ir})$) that could potentially correlated with the government attitudes towards city lockdown, then our estimation will be biased. In order to solve this potential endogeneity problem, we take advantage of the fact that the outbreak of the virus in Wuhan but not other cities is exogenous, and cities' lockdown policies reflect the city characteristics, unobserved government attitude, but also mainly determined by the degree of virus exposure, measured by the amount of people travel from Wuhan before the city lockdown and the distance to Wuhan, as studied in Fang et.al (2020). In Appendix Table **??**, we show the larger the accumulated travel flow from Wuhan 20 days before the Wuhan's lockdown, the larger the negative demand shock each city experienced at the end of March. Hence we use utilize these two variables as the instrumental variable. In other words, we estimate the impact from only the exogenous part of LMS that aroused only by the interactions with Wuhan.

6 Main results

6.1 Robustness check with various model setups

We first pool all people who migrated to cities in 2019 together, and estimate the overall reallocation elasticity ($(Pr(s_t^* = 3|4))$ and unemployment elasticity ($((Pr(s_t^* = 1)))$ with various sample selections (Top 45 cities only and all cities) and measurement of labor demand shocks, in order to show the robustness of our results. In Table 2, we report the marginal effect of the local labor demand shocks in the previously working location ($\Delta P_{r_{t-1}^*} / P_{r_{t-1}^*}$, with the predicted probability of reallocation . Hence according to the Equ.4, we can calculate the short-term elasticity. We first show the baseline model with only the labor demand shock measurement (Column 1,3,5,7), and then add more control variables (Column 2,4,6,8) to study how other factors affect people's reallocation decisions. Column 1-4 constrains with the Top 45 cities and Column 5-8 with all cities. For each sample selection, we utilize the LMS definition in Equ.7 in the first two columns (Column 1,2,5,6) and the definition in Equ.8 in the last two columns (Column 3,4,7,8). Panel A reports the unemployment elasticity and Panel B the reallocation elasticity. As we can see, the cities' labor demand shock significantly affects the possibility of people reallocate and stay unemployed, with the labor demand shock worsens by 1%, individuals' possibility of unemployment will increase by around 13.3-14.1 percentage points, given the predicted y as 0.353-0.359, this means that the elasticity of unemployment to the labor demand shock is within the range from -0.37 to -0.39, with the Top 45 cities sample gives slightly lower elasticity but not significant and the two measurement also delivers very similar results. It's worthy to notice that the sample size is also comparable in the two sample selection, suggesting that most of the out-migration go to the large cities. For reallocation, with the labor demand shock worsens by 1%, individuals' possibility of reallocation will increase by around 5.6-6.8 percentage points, given the predicted y as 0.321-0.323, this means that the elasticity of reallocation to the labor demand shock is within the range from -0.17 to -0.21, which is also consistent with different labor market demand measurement, different sample selection, with and without control variables. The magnitude of reallocation is much smaller than the unemployment, suggesting relatively higher inertia of spatial mobility for the low-skilled population in general.

To furthermore establish the robustness of our results, we first conduct a placebo test in which we run a Probit regression model of the LMS in 2020 on the the transition to unemployment from 2018 to 2019. If the observed unemployment in Table 2 indeed comes from the labor demand shock in early 2020, the transition to unemployment from 2018 to 2019, should not be significantly affected by this LMS. Panel of in Table 3 exactly follow our prediction, with the marginal effect insignificantly and almost zero (around 0.002-0.006). The probability of transition to unemployment for this group of people is also small in 2019 (the predicted value is around 0.021-0.03). Another way to show robustness is to conduct IV-Probit regression which takes into consideration of the potential correlation of individuals' unobserved migration preferences and city lockdown policies. As we can see in Panel B of Table3, IV-Probit regressions gives very similar results compared to the Probit model, suggesting the robustness of our finding and that the exogeneity assumption is valid in our setup. Hence in the following analysis, to guarantee the estimation efficiency, we'll only consider the Probit model.

To explore the robustness of our results and solve the potential serial correlations in the error terms, inspired by Bertrand et al. (2004) and Rosenbaum (2002, 2009) who argue that randomization tests have good properties in a variety of non-experimental settings, we implemented a randomization test based on placebo laws. We randomly assign the LMS between -1 to 1 to each individual, estimated the regression model using the placebo treatment instead of the real one, and repeated 2000 times to build a distribution of placebo effects. Theoretically, these placebo laws should have no effect on the outcome, with only seldom resulting in large effects by chance. Fig. 8 plots kernel density estimates of the placebo law effects. The distributions vary somewhat across the three regression setups; however the means are approximately zero, which implies that the effect estimator is unbiased. The effect estimate from the actual data - -0.141 - falls in the extreme left tail of the distribution of placebo effects, which suggests it is unlikely that the effect was observed due to chance. Based on the placebo law distribution, the probability of observing an effect of -0.141 when the true effect is 0 is well below .01. Thus, the placebo laws tests cast doubt on the possibility that our conclusions are invalid because of biased standard errors, since this test does not require assumptions about clustering and serial correlation but only that the treatment variable is independent of the potential outcomes.

	Top 45 cities				All cities			
	$\Delta D_1\%$		$\Delta D_2\%$		$\Delta D_1\%$		$\Delta D_2\%$	
Panel A: $k_{it,j}^{j*} = 0$								
LMS	-0.136***	-0.136***	-0.133***	-0.133***	-0.141***	-0.141***	-0.139***	-0.138***
	(0.024)	(0.024)	(0.025)	(0.025)	(0.023)	(0.023)	(0.024)	(0.024)
Predicted $s^* = 2$	0.355	0.353	0.355	0.353	0.361	0.358	0.361	0.359
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	9191	9116	9191	9116	10585	10503	10585	10503
Panel B: $k_{it,j}^{j*} \in [1,\ldots,I+1] \neq k_{i,t-1}^{j*}$								
LMS	-0.068***	-0.062***	-0.065***	-0.058***	-0.066***	-0.059***	-0.063***	-0.056***
	(0.02)	(0.019)	(0.02)	(0.02)	(0.019)	(0.019)	(0.02)	(0.02)
Predicted $s^* = 3\|4$	0.322	0.321	0.322	0.321	0.323	0.322	0.323	0.322
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	10242	10162	10242	10162	11789	11700	11789	11700

[a] All probit regression are run with robust standard errors clusterd by village. Standard errors reported in the parenthesis, with *** represents significance level ¡1%, ** for 5% and * for 10%
[b] We report the marginal effect for the two panels, with the predicted y also reported.

Table 2: Main results for Probit models

	$\Delta D_1\%$				$\Delta D_2\%$			
	Probit		IV-Probit		Probit		IV-Probit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: $k^{j*}_{it,j} = 0$ **in** $t = 2019$								
LMS	0.005	0.005	0.006	0.006	0.002	0.003	0.004	0.004
	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)
Predicted *reallocate* $= 1$	0.029	0.021	0.029	0.021	0.03	0.022	0.03	0.022
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	10196	9854	10196	9854	11872	11468	11872	11468
Panel B: IV-Probit results for $k^{j*}_{it,j} = 0$								
LMS	-0.122***	-0.116***	-0.12***	-0.112***	-0.112***	-0.111***	-0.11***	-0.108**
	(0.023)	(0.037)	(0.024)	(0.039)	(0.015)	(0.042)	(0.016)	(0.043)
Predicted *reallocate* $= 1$	0.355	0.353	0.355	0.353	0.361	0.359	0.361	0.359
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	9125	9050	9125	9050	10424	10342	10424	10342

[a] All probit regression are run with robust standard errors clusterd by village. Standard errors reported in the parenthesis, with *** represents significance level ¡1%, ** for 5% and * for 10%
[b] We report the marginal effect for the two panels, with the predicted y also reported.

Table 3: Main results for robustness checks

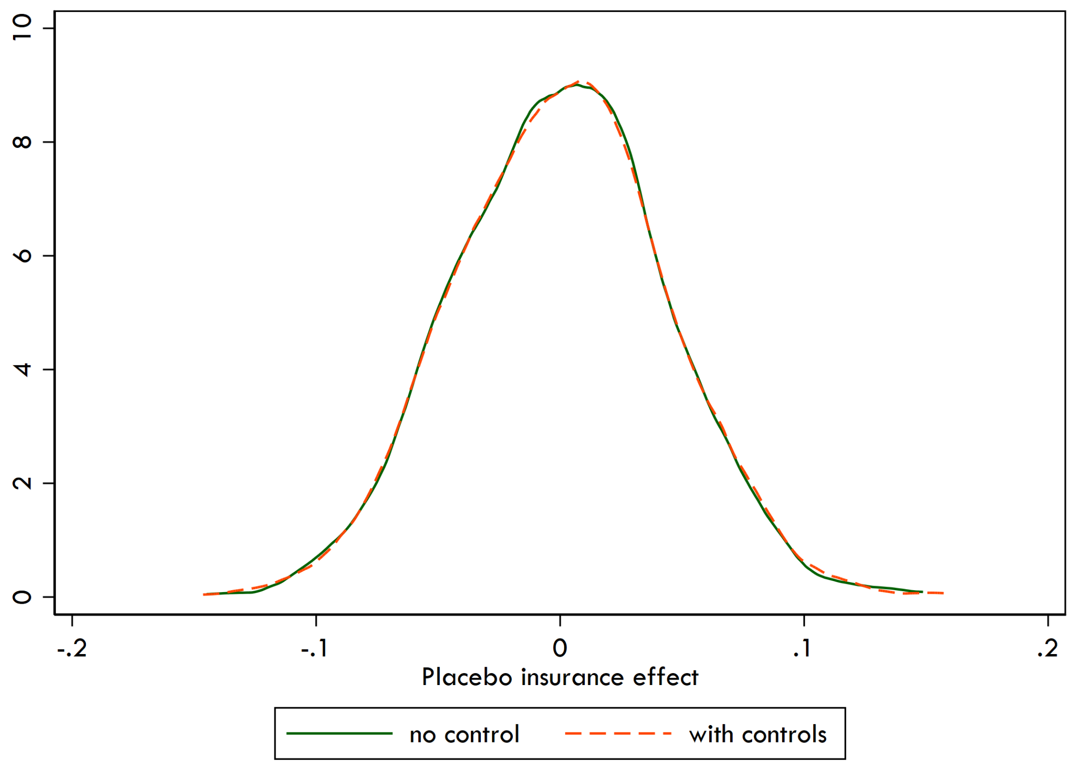


Figure 8: The graph shows a kernel density plot of the distribution of 2000 placebo estimates of the effects of demand shocks on possibility of staying in unemployment. Each line shows the distribution from different regression setup. The means of the three distributions are 0.000015 and -0.00003 respectively. Hence the estimate from the actual data, which is -0.141, is extremely unlikely to have arisen by chance.

6.2 Heterogeneity by age and gender

We then separate the individuals into different types and study their behavior patterns respectively. We first consider the age and gender differences, we separate the population into (1) young and middle-aged male (age between 16-45); (2) young and middle-aged female (age between 16-45); (3) old male (age larger than 45); (4) old female (age larger than 45). As we stated in the last subsection, given the fact that various model setup gives similar results, for the following analysis, we only show the results based on all cities with LMS measured by Equ.7[19]. We report the results in Table 4.

We notice that the males behave significantly different from the females. In terms of unemployment, the young and middle-aged males experience the lowest possibility of unemployment(29.7%) and males in general have larger trend of reallocation (33.6%-34.7%). While the females have higher unemployment rate, with 44$ of the old females turn to unemployment in early 2020, and smaller possibility of transition. For the young and middle-aged males, with the LMS increased by 1%, the possibility of $k_{it}^{*} = 0$ increases by 13 points and the possibility of reallocation increases by 11.1 percentage points, with the elasticity of unemployment around 0.448 and the elasticity of reallocation around 0.330. Even though the reallocation elasticity is still lower than the unemployment, but it's much larger than the average reallocation elasticity in Table 2 (0.204). While for the females in the same age, the elasticity of unemployment is similar (0.392) but for the old females, the unemployment response is insignificant, which means that the old females experience large possibility of unemployment (0.44) from all cities. One of the possible explanation for this, is that the young and middle-aged females are almost perfect substitutes for the old females, then the old females also experience larger unemployment even the overall labor demand shock is not large.

In order to dig into the details of individuals' heterogeneous response, we then explore the multinomial Probit Model setup. Among the people who used to work outside of Xin County in 2019 and already get employed by the end of 2020, March, we study their location switching behavior in response to the various demand shock. With the labor demand shock in the destination city, among the people who can find jobs, they can either switch to another city with larger uncertainty, or staying within the County in which they have more social connection but lower income level. In particular, we are interested in the gender differences in terms of responsiveness and response type.

From the regression results plotted in Figure 9: we can see that among the people who reallocate, females and males show completely different patterns. Compared to transiting back to hometown, more males tend to switch to another city outside of the county for higher income but also higher uncertainty, while the females are more tend to stay in the hometown, with less income but more security.

We also explore the weekly employment data, to examine the timing of individuals'

[19]For the results based on other measurements, please contact the author

	$16 \leq age \leq 45$				$age > 45$			
	Males		Females		Males		Females	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: $k^{j*}_{it,j} = 0$								
LMS	-0.137***	-0.134***	-0.152***	-0.143***	-0.153***	-0.148***	-0.078*	-0.075
	(0.031)	(0.031)	(0.045)	(0.045)	(0.034)	(0.033)	(0.047)	(0.048)
Predicted $s^* = 2$	0.297	0.299	0.388	0.385	0.379	0.377	0.44	0.438
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	3574	3606	1728	1718	3944	3912	1293	1285
Panel B: $k^{j*}_{it,j} \in [1, \ldots, I+1] \neq k^{j*}_{i,t-1}$								
LMS	-0.111***	0-.111***	-0.033	-0.025	-0.02	-0.023	-0.062	-0.052
	(0.029)	(0.029)	(0.038)	(0.037)	(0.03)	(0.031)	(0.042)	(0.042)
Predicted $s^* = 3\|4$	0.336	0.336	0.257	0.259	0.344	0.347	0.271	0.274
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	4071	4036	1959	1970	4255	4291	1440	1408

[a] All probit regressions are run with robust standard errors clustered by village. Standard errors reported in the parenthesis, with *** represents significance level ¡1%, ** for 5% and * for 10%
[b] We report the marginal effect for the two panels, with the predicted y also reported.

Table 4: Probit results by gender and age group

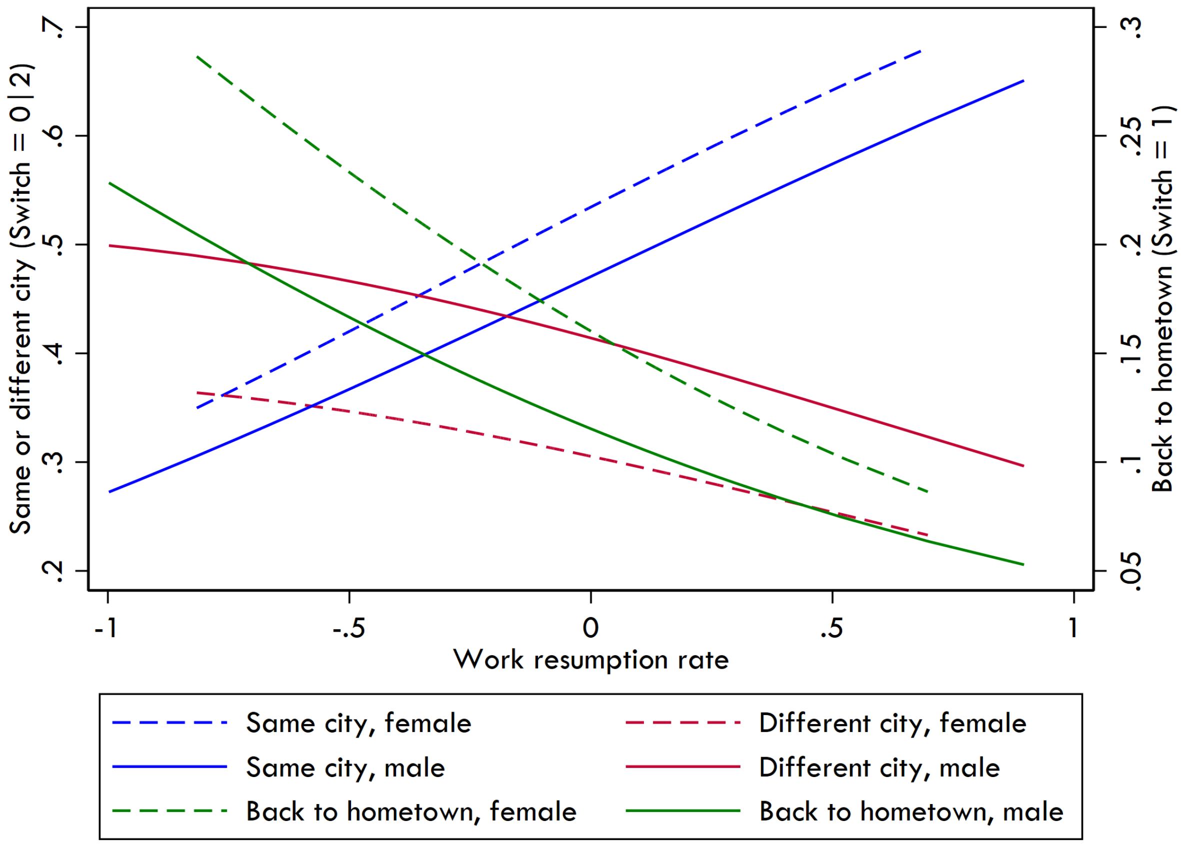


Figure 9: Location switching behavior between females and males

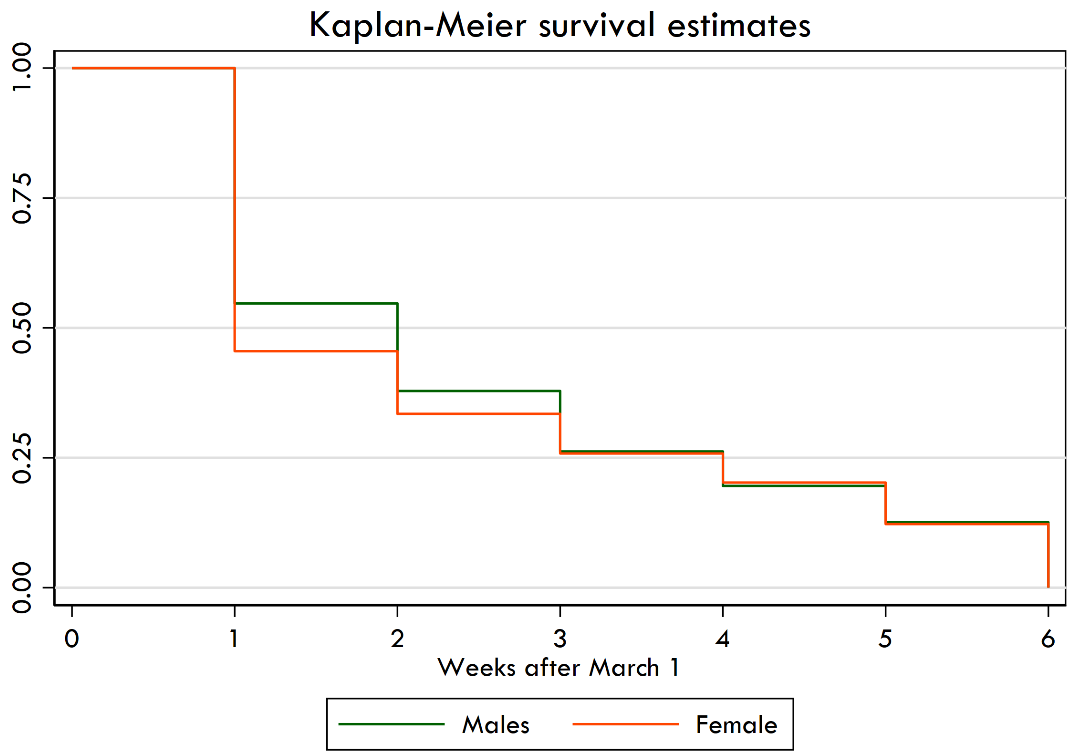


Figure 10: Survival analysis by gender

responses. We employ Cox(1972)'s Discrete Survival function 12:

$$\lambda(t_j|X_i) = \lambda_0(t_j)\exp\{X'\beta\} \tag{12}$$

where $\lambda(t_j|x_i)$ is the hazard rate at time t_j as the conditional probability of event (getting employed) at that time given that one has survived (stay unemployed) to that point that for an individual with covariate values X_i. The covariate includes the labor market shocks $\text{LMS}_{i,t}$ and other demographic variables. According to the survival estimates depicted in Figure 10, the females started work earlier, with significant higher hazard rate (probability of getting employed) at 1.05 (std = 0.026).

Combining these two sets of results together, we can see that the negative demand shocks drive people to adjust (either reallocate or stay in local) in different ways, with females adjust earlier but more to the local areas with pecuniary loss, and males to another cities with higher uncertainty. Even with large heterogeneity, we can still see that among the low-skilled workers, the unemployment elasticity are all higher (around 0.4) than reallocation, suggesting inertia for spatial mobility for all types of individuals.

6.3 Heterogeneity by education level

Current literature suggests people with higher education are more flexible in terms of spatial mobility and reallocation. Though we focus on the individuals from low-income families and the education level and human capital is relatively lower than the rest of the people. We can still dig into the variation in education and classify individuals by education level, and explore how education will affect people's responsiveness to demand shocks. Table **??** report the separate regression results for the four education levels: elementary and below, middle school, high school and 3-year college and above. Given the fact that the education level is highly correlated with ages, caused by the education resource expanding in the past 50 years, we limit to young and middle-age people (16-45 year old) in our regressions to isolate the education effect.

As we can see, for people under 45, most of people have middle school education. For people with only elementary education, the predicted value for $k^*_{it} = 0$,i.e. stay unemployed is 42.6%. The negative shock worsens by 1%, the possibility of staying in unemployment increases by 0.229 percentage points, hence the corresponding elasticity is -0.538, while that for people with college education above is -0.43 with the lowest unemployment rate at 28.6%. Besides, the unemployment among people with middle school education, high school education and 3-year college is comparable, with the chi-test statistics = 0.02. Even though the people with only elementary education indeed has higher level of inertia with respect to the demand shock, the education improvement above middle school cannot significantly lower the inertia, suggesting that among the rural migrants, the value of higher education in terms of job hunting is quite limited. Similarly we conduct multinomial Logit model to study the direction of reallocation. We group people into high school above and below, with results shown in Figure 11. We can see that among the people who reallocate, people with different education levels show completely different patterns. Compared to transiting back to hometown, more high educated people tend to switch to another city outside of the county for higher income but also higher uncertainty, while the low educated people are more tend to stay in the hometown.

7 Mechanism

In the previous section, we provide evidence suggesting high inertia of reallocation within these individuals coming from the most disvantaged low-income families in rural China. Besides, we show that females compared to males, are more reluctant to reallocate, especially to new cities. They are more tend to bear the pecuniary loss and quickly settle down with low salary job in the local area. Here in this section, we'll try to explain this high inertia for these people, we'll argue, it's the dependence on social network, that stops those people to reallocate.

7.1 Dependence on social network

In almost all migration literature using Indian data, the community-based social network is defined within caste lines. This definition is based on the fact that there are frequent social interactions and close ties within the caste, which consists of thousands of households and spans a wide area covering many villages, support very connected and exceptionally

	Elementary		middle		high		college	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: $k^{j*}_{it,j} = 0$								
LMS	-0.229**	-0.22**	-0.124***	-0.132***	-0.131**	-0.132**	-0.13***	-0.124***
	(0.094)	(0.095)	(0.035)	(0.036)	(0.054)	(0.053)	(0.046)	(0.046)
Predicted $s^* = 2$	0.425	0.426	0.335	0.335	0.313	0.313	0.288	0.286
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	391	391	2714	2714	1005	1005	1182	1182
Panel B: $k^{j*}_{it,j} \in [1, \ldots, I+1] \neq k^{j*}_{i,t-1}$								
LMS	-0.084	-0.084	-0.084***	-0.082**	-0.097*	-0.094*	-0.095**	-0.092**
	(0.075)	(0.076)	(0.032)	(0.032)	(0.056)	(0.055)	(0.044)	(0.043)
Predicted $s^* = 3\|4$	0.309	0.311	0.311	0.312	0.328	0.328	0.296	0.296
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	434	434	3105	3105	1126	1126	1330	1330

[a] All probit regressions are run with robust standard errors clustered by village. Standard errors reported in the parenthesis, with *** represents significance level ¡1%, ** for 5% and * for 10%
[b] We report the marginal effect for the two panels, with the predicted y also reported.

Table 5: Probit results by education level

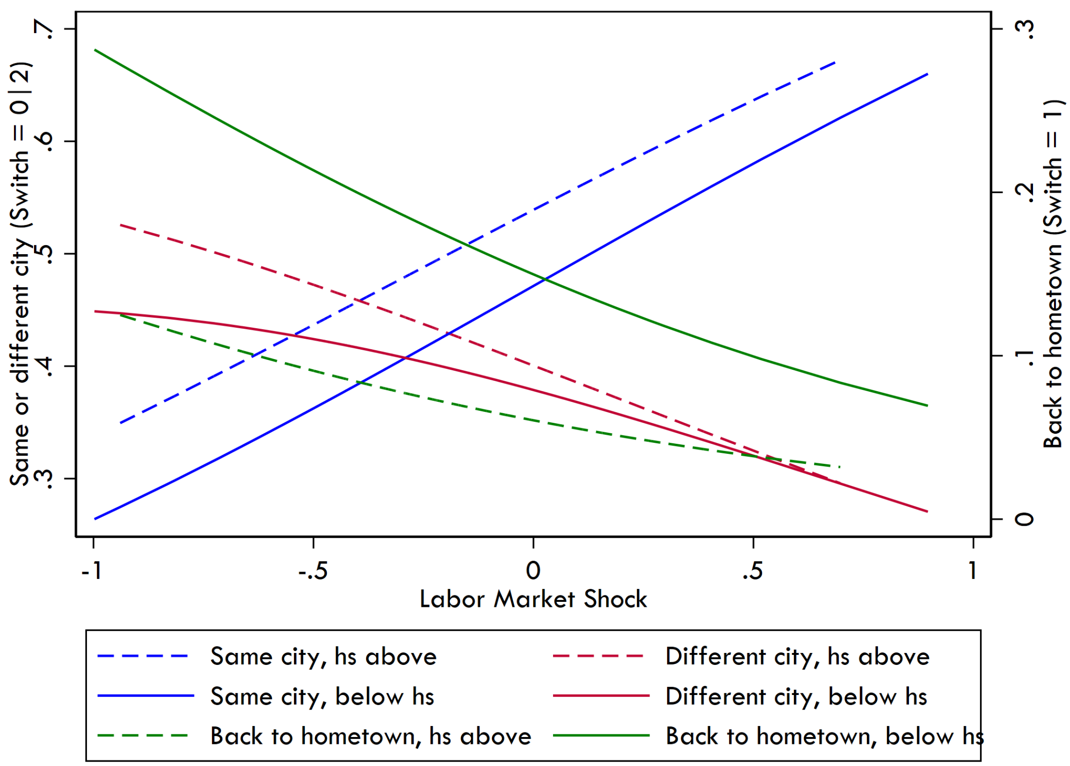


Figure 11: Location switching behavior between high and low education level

extensive insurance networks (Caldwell, Reddy, and Caldwell 1986; Mazzocco and Saini 2012). In China, we instead define the network by village, which is the rural community in China.

As we state in the background section, that the most of the rural migrants undertake low-skills job with little contract or insurance protection, and also high substitutable and unstable. In the survey, we asked how they obtain their current job. Figure 12 shows the distribution of the job obtaining channels. As we can see, over 85% of poor rural workers got their jobs by social networks, either through family links, or village links or workmates. The high dependence on the social network makes them reluctant to leave where they used to work, also harder to find new jobs in unfamiliar places.

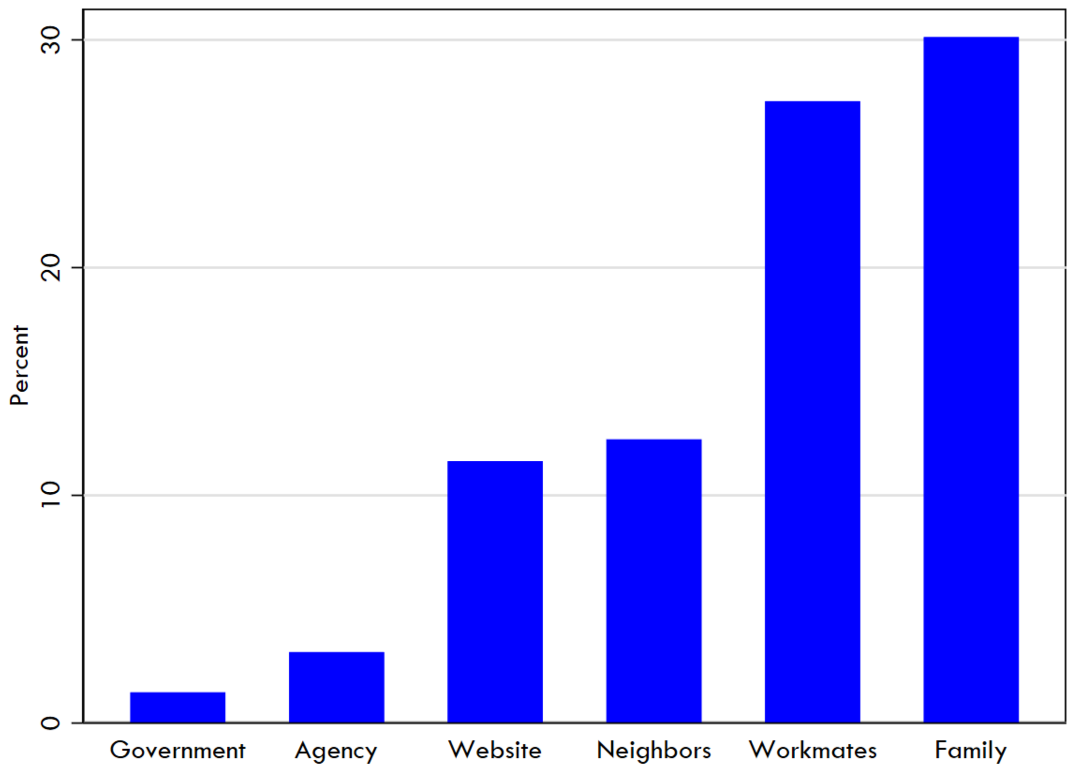
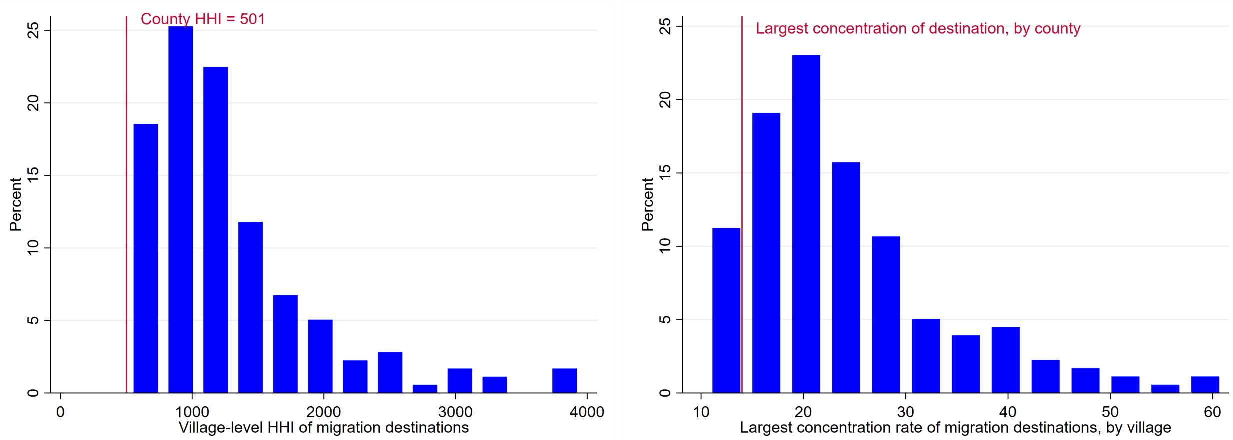


Figure 12: Distributions of job hunting channels

In Munshi and Rosenzweig (2016), they argue that the rural migrants tend to migrate as a group to utilize the previous community-based social network. In order to directly test people's network dependence and the inertia to response and inspired by Mushi and Rosenzweigh(2016), we use the individual's working location data to calculate the concentration of migration destinations within the village,, we construct measurements for people's network dependence. As we can see in Figure 12, the biggest channel for job hunting is through families and villagers (42.56% combined), hence we should observe high similarity of migration destinations within village. Given the next biggest channel is through workmates in the destination, this similarity should be strengthen along the time.

In other words, the higher the village similarity in destination, the higher dependence the villagers will be on the social network in terms of job hunting. We utilize individual's working location in 2019, and calculate the HHI[20] and CR1[21] of migration locations within the village. Figure 13 show the distribution of village-level HHI and CR1 in Xin County, with the county-level HHI and CR1 highlighted for comparison. As we can see, there is a large variation in terms of destination concentration rate, with 43% villages have HHI lower than 1000, and a few above 3000, 70% village CR1 larger than 20% (means over 20% of all migrants work in the same city) and the maximum up to 60% (over 60% of all village migrants in one city). This variation in destination concentration provides us a chance to directly test whether the social network dependence can explain the rural migrants' inertia to responses.



((a)) HHI

((b)) CR1

Figure 13: Migration destination concentration rate, 2019

7.2 Network dependence and inertia to reallocate

On top of the baseline logit model we considered in the last section, we add in a interaction term Network$_{i,t}$ and LMS in the covariates X_i. If the coefficient of LMS is β and the coefficient of this interaction term is α, then the overall effect of LMS is $\beta \times (1+\alpha)$. Given the fact that $\beta < 0$, then if as we predicted, i.e., the tighter the social ties, the larger the impact of LMS on indiviudals' behavior. Then α for unemployment regression should be positive,

[20] $HHI = \sum_{i=1}^{N} city_i^2$, in which $city_i$ is the percentage of people migrate to $city_i$ in 2019

[21] $HHI = \max_{i=1}^{N} city_i$, in which $city_i$ is the percentage of people migrate to $city_i$ in 2019

suggesting the larger the negative demand shock, with larger social network dependence, the possibility of unemployment which is $|\beta \times (1+\alpha)|$ is getting larger. While the $alpha$ for reallocation should be negative, hence the reallocation possibility $|\beta \times (1+\alpha)| < |\beta|$ will be smaller

In Table **??** we report the regression results, and in **??**, we report the marginal effect. For simipilicity for interpretation, we normalize the HHI as:

$$HHI_{norm} = \frac{HHI - \overline{HHI}}{std(HHI)} \tag{13}$$

As we can see, the community-based social connections significantly affects the degree of inertia and differently for each gender. For unemployment, when the connection gets stronger by 1 std.dev, the margin effect of labor demand shock increase by 0.05 percentage points for males, which is 1/3 of the LMS effect, and around 0.08 percentage points, which is 2/3 of the LMS effect. As we can see, the females depend on the social connection more than the male counterparts. For the reallocation, the higher the social connection, the higher inertia for reallocation. As we can see, all interaction coefficients are negative. However, only for females, the coefficient is significant, with LMS itself insignificant for most setups. This suggest, the social network strongly affects the females' possibility to reallocate, or in other words, without social connections in the new place (or high concentration of people in previous location), the females will not tend to reallocate no matter what the market condition is.

8 Conclusion

In the beginning of 2020, Chinese cities autonomously adopted various levels of city lockdown in response to the sudden outbreak of COVID-19. Starting from the end of February, along with the virus under control, cities started a resumption of work with different timing, hence imposed different demand shocks to rural migrants who used to work there. In this paper, we utilize city lockdown policies as an exogenous quasi-experiment and examine rural migrants' degree of responsiveness to the sudden demand shock, and we also explore the mechanism behind this phenomenon.

We find that the poor population is highly affected because of the high level of stigma of their work location choices. The poor population tend to work in the same location and are reluctant to switch to other cities even with the original cities locked down. With the work resumption rate of the person's previous working location drops 1%, the person's unemployment chance increases 0.4%. The survival analysis shows that even for the people who get employed eventually, the large labor demand shock of the previous working location significantly prolongs the unemployment duration.

Furthermore, we look into the mechanism behind this phenomenon, and find this high level of stigma can largely be explained by the dependence on the informal insurance

	Males				Females			
	$\Delta D_1\%$		$\Delta D_2\%$		$\Delta D_1\%$		$\Delta D_2\%$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: $k^{j*}_{it,j} = 0$								
LMS	-0.152***	-0.144***	-0.147***	-0.138***	-0.12***	-0.116***	-0.119***	-0.114***
	(0.019)	(0.019)	(0.019)	(0.02)	(0.03)	(0.03)	(0.031)	(.032)
$LMS \times HHI$	0.049***	0.055***	0.047***	0.053***	0.079***	0.085***	0.076***	0.082***
	(0.015)	(0.015)	(0.013)	(0.013)	(0.026)	(0.026)	(0.022)	(0.022)
Predicted $s^* = 2$	0.34	0.339	0.341	0.339	0.41	0.408	0.41	0.408
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	7563	7496	7563	7496	3022	3004	3022	3004
Panel B: $k^{j*}_{it,j} \in [1, \ldots, I+1] \neq k^{j*}_{i,t-1}$								
LMS	-0.074***	-0.073***	-0.07***	-0.069***	-0.043*	-0.043	-0.043	-0.043
	(0.018)	(0.018)	(0.018)	(0.019)	(0.026)	(0.026)	(0.027)	(0.027)
$LMS \times HHI$	-0.013	-0.017	-0.013	-0.016	-0.055***	-0.06***	-0.048***	-0.053***
	(0.014)	(0.014)	(0.011)	(0.012)	(0.021)	(0.021)	(0.017)	(0.017)
Predicted $s^* = 3\|4$	0.341	0.34	0.341	0.34	0.264	0.263	0.264	0.262
Control	No	Yes	No	Yes	No	Yes	No	Yes
obs	8378	8305	8378	8305	3411	3392	3411	3392

[a] All probit regressions are run with robust standard errors clustered by village. Standard errors reported in the parenthesis, with *** represents significance level ¡1%, ** for 5% and * for 10%
[b] We report the marginal effect for the two panels, with the predicted y also reported.

Table 6: Probit results by gender and age group, Social network effect

provided by the village social network. We use the concentration rate of working location within the village as an index for the strength of social network within the village, and find that when the network gets stronger by 1 std.err, the margin effect of work resumption increase by 0.09%, which is 15% of the average effect. This means that the stronger the social network one person has, the more reluctant this person is to switch to other working locations, even if the other location means higher chances of getting employed.

To sum up, our study shows as to the working choices, if the loss in network insurance is sufficiently severe, then higher paying job opportunities will go unexploited. The high depends on social networks induces significant inertia of reallocation due to the regional labor demand shocks. This study has strong implications in regards to the policies with respect to the rural labor force. For instance, in the beginning of work resumption, some cities establish collaboration with poor counties and attempt to hire workers directly through government advertisement but ended up with few people applied. This can be explained by the channel proved by our study that people refuse to switch to unfamiliar locations by themselves with the loss of the established social network. On the other hand, the perfection of formal insurance by the government safety net can substitute the informal insurance and enhance the mobility of those low-skilled workers and attenuate the misallocation.

References

T. Allen and C. Arkolakis. Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

N. Angelov, D. Waldenström, et al. The impact of covid-19 on economic activity: Evidence from administrative tax registers. Technical report, 2021.

G. Bryan, S. Chowdhury, and A. M. Mobarak. Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica*, 82(5):1671–1748, 2014.

B. C. Cadena and B. K. Kovak. Immigrants equilibrate local labor markets: Evidence from the great recession. *American Economic Journal: Applied Economics*, 8(1):257–90, 2016.

D. Coen-Pirani. Understanding gross worker flows across us states. *Journal of Monetary Economics*, 57(7):769–784, 2010.

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, 2016.

R. Dix-Carneiro and B. K. Kovak. Trade liberalization and the skill premium: A local labor markets approach. *American Economic Review*, 105(5):551–57, 2015.

D. Donaldson and R. Hornbeck. Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016.

J. Kennan and J. R. Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011.

M. D. Kugler, M. Viollaz, D. Vasconcellos Archer Duque, I. Gaddis, D. L. Newhouse, A. Palacios-Lopez, M. Weber, et al. How did the covid-19 crisis affect different types of workers in the developing world? Technical report, The World Bank, 2021.

R. Molloy, C. L. Smith, and A. Wozniak. Internal migration in the united states. *Journal of Economic perspectives*, 25(3):173–96, 2011.

E. Moretti. *The new geography of jobs*. Houghton Mifflin Harcourt, 2012.

M. Morten. Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy*, 127(1):1–46, 2019.

M. Morten and J. Oliveira. Paving the way to development: Costly migration and labor market integration. Technical report, National Bureau of Economic Research, 2016.

K. Munshi and M. Rosenzweig. Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98, 2016.

J. Roback. Wages, rents, and the quality of life. *Journal of political Economy*, 90(6):1257–1278, 1982.

H. S. Rosen. Housing decisions and the us income tax: An econometric analysis. *Journal of Public Economics*, 11(1):1–23, 1979.

Y. Su, Y. Hua, and X. Liang. Toward job or amenity?: Evaluating the locational choice of internal migrants in china. *International Regional Science Review*, 42(5-6):400–430, 2019.

T. Tombe and X. Zhu. Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–72, 2019.

A. Wozniak. Are college graduates more responsive to distant labor market opportunities? *Journal of Human Resources*, 45(4):944–970, 2010.

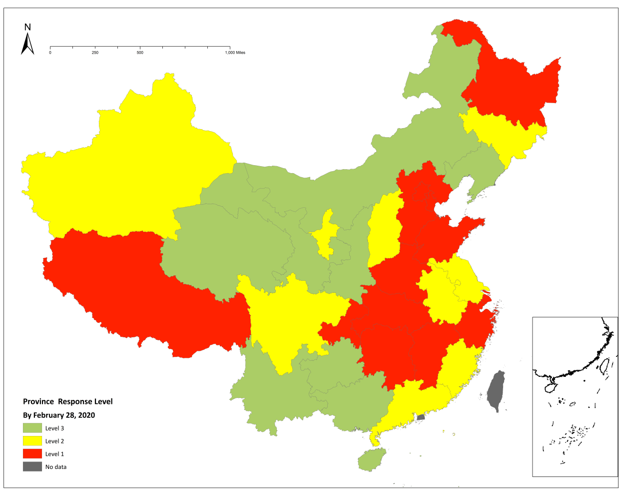
9 Appendix

Variable	obs	mean	std.dev	min	max
Sample selection					
Work in big cities in 2019	10,968	0.838	0.369	0	1
$\Delta_D\% \in (-1,1)$	10,968	0.949	0.22	0	1
Individual choices					
unemploy = 1	10,483	0.33	0.47	0	1
adjust = 1	10,968	0.362	0.481	0	1
reallocate = 1	10,968	0.264	0.441	0	1
local = 1	10,968	0.098	0.298	0	1
Individual characteristics					
elementary degree	10,878	0.214	0.41	0	1
middle school degree	10,878	0.56	0.496	0	1
high school degree	10,878	0.115	0.318	0	1
3-year college and above	10,878	0.111	0.314	0	1
age	10,968	42.228	12.391	17	78
Male dummy	10,968	0.716	0.451	0	1
Family demographics					
# of children under 5	10,968	0.143	0.398	0	3
# of students	10,968	0.497	0.735	0	4
# of adult males	10,968	1.495	0.751	0	5
# of adult females	10,968	1.16	0.855	0	4
per capita income (thousand)	10,968	9.49	4.133	2.369	121.648

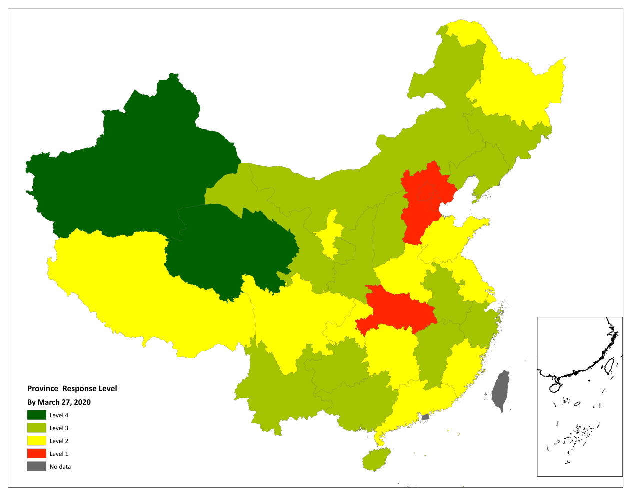
Table 7: Descriptive Statistics

Policy	Beijing	Shenzhen
Prevention Level down	April 30	Feb 24
Quarantine requirement	All quarantine for 14 days	No quarantine if from low-risk areas
Public transportation	Capacity control at 65% till April 30	back to normal
Stimulation policies	Rent exemption or 50% reduction only for February	18 policies including rent exemption, social insurance payment exemption etc.

Table 8: Policy comparison between Beijing and Shenzhen, till March 30, 2020

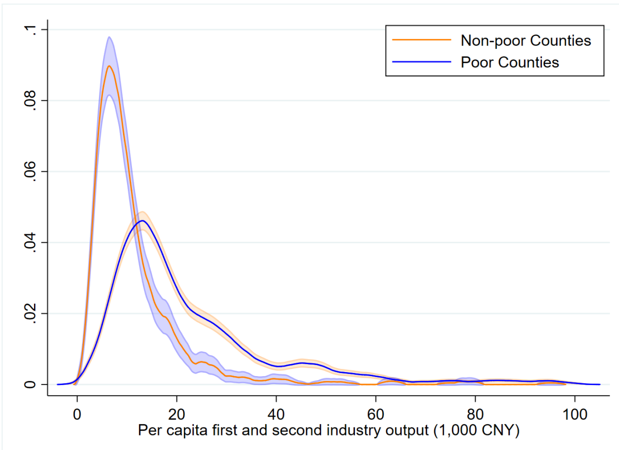


((a)) Province response level by February

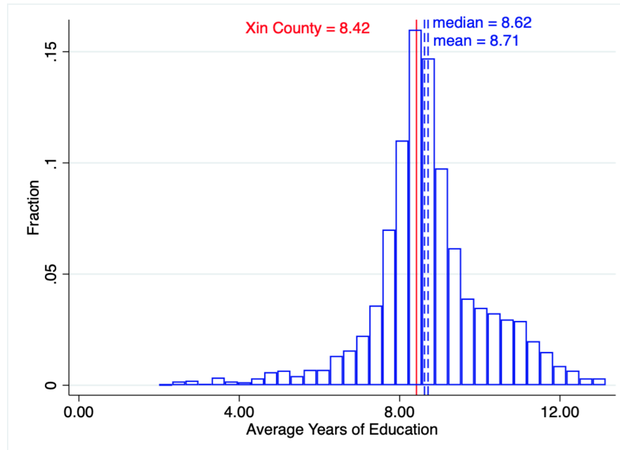


((b)) Province response level by March

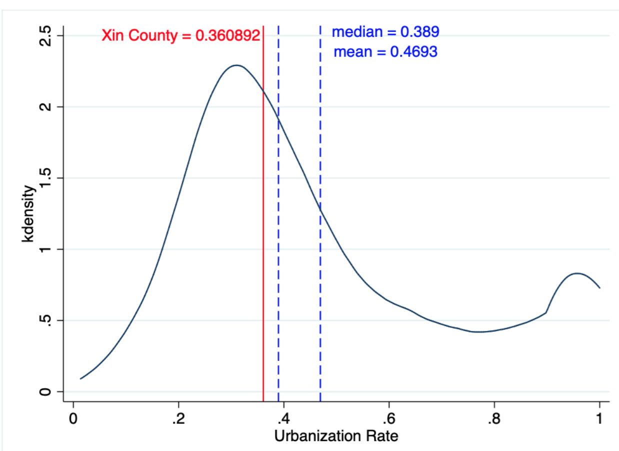
Figure 14: Spatial distribution of provincial response level



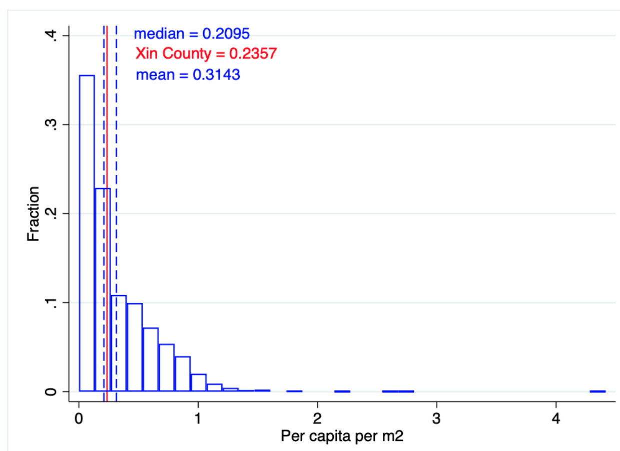
((a)) Per capita GDP(1k Yuan)



((b)) Average Years of Education

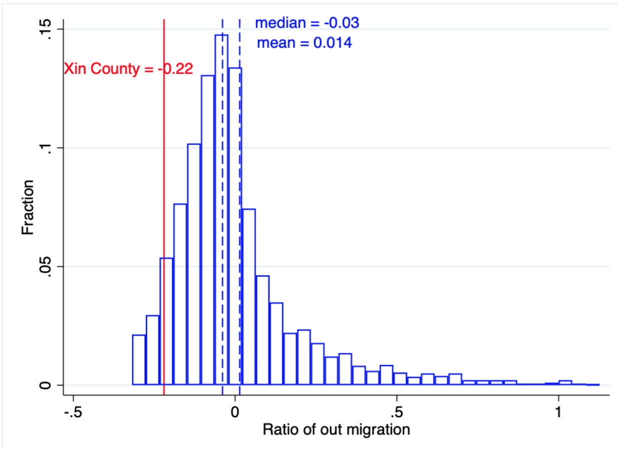


((c)) Urbanization Rate

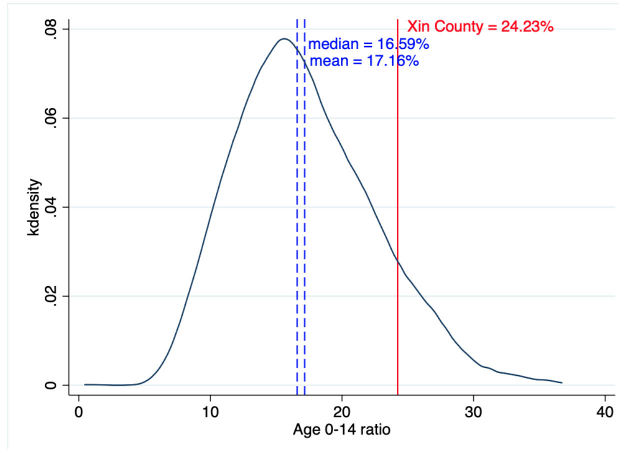


((d)) Population Density

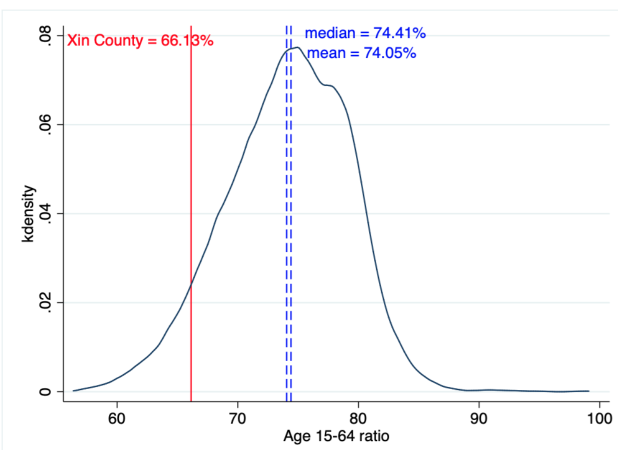
Figure 15: Distributions of Socioeconomic Variables



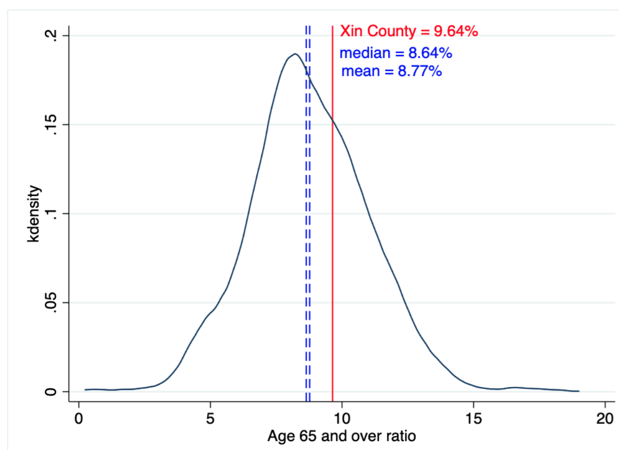
((a)) Ratio of out-migrating people



((b)) Ratio of people under 14

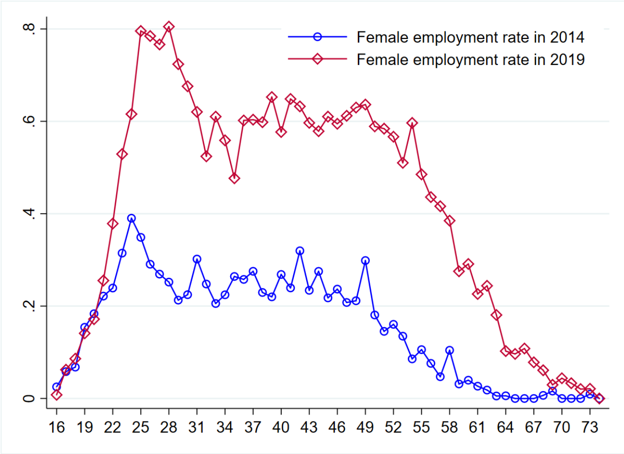


((c)) Ratio of people 15-64

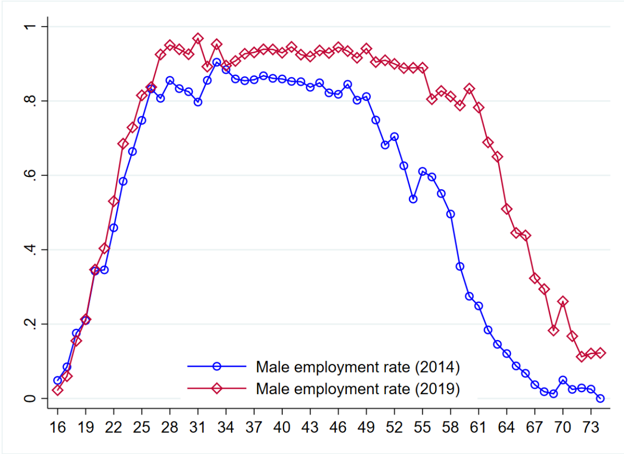


((d)) Ratio of people over 65

Figure 16: Distributions of Demographic Variables Related to Out-migration



((a)) Females



((b)) Males

Figure 17: Labor participation rate of females and males in Xin County, 2014 vs 2019

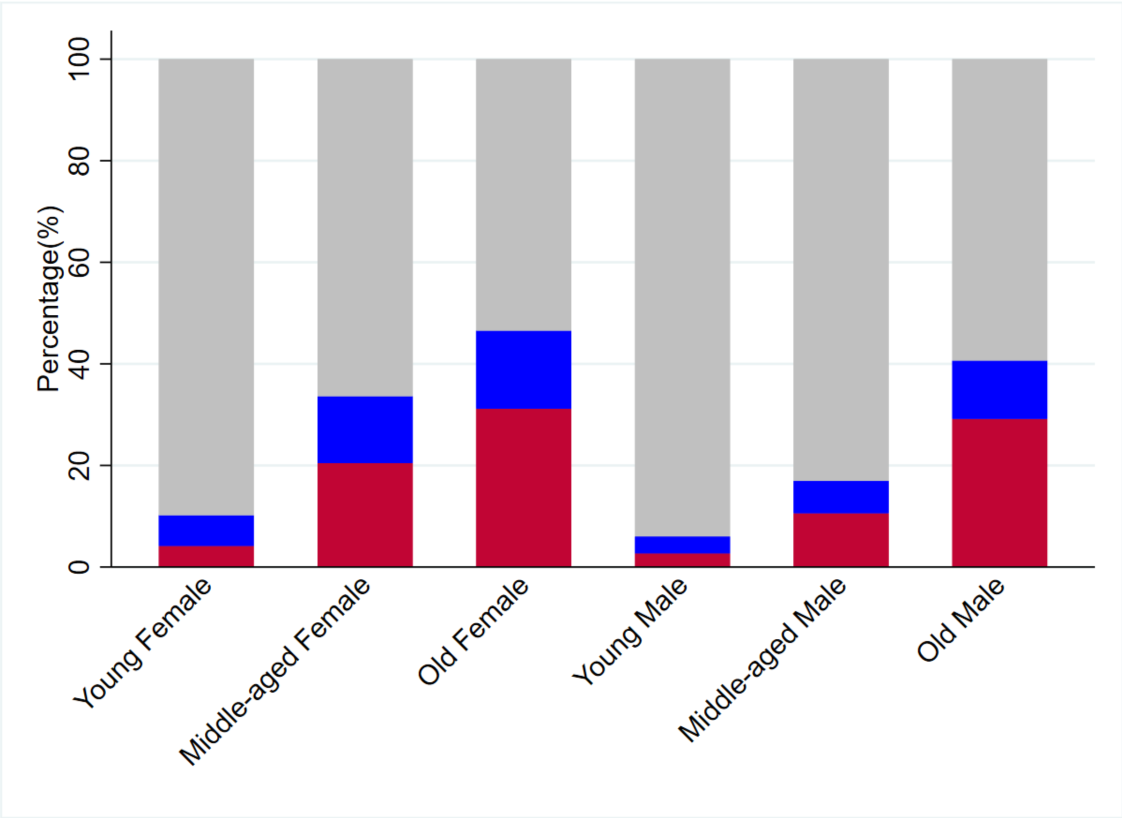


Figure 18: Working locations by gender and age group, in 2019

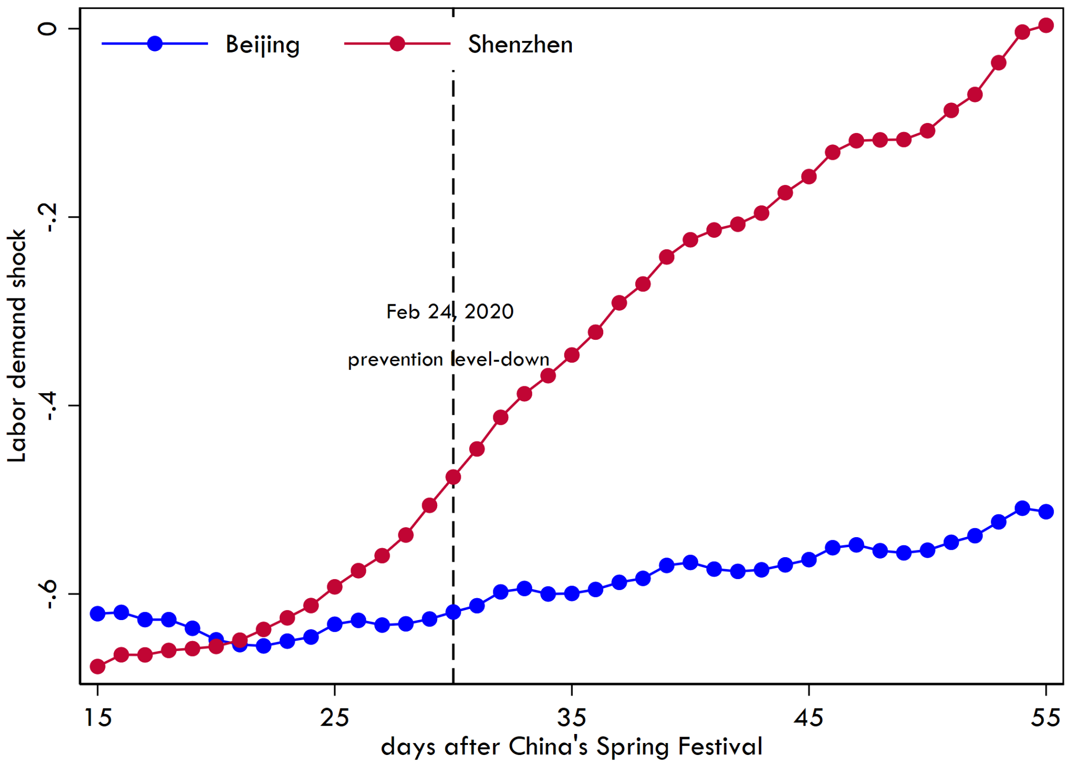


Figure 19: Comparison of time trends of labor demand shocks in Beijing and Shenzhen, 15 days - 55 days after the Spring Festival