# COVID-19 Pandemic and Unemployment: Evidence from Mobile Phone Data in China

Teng Li, Panle Jia Barwick, Yongheng Deng, Xinfei Huang, Shanjun Li*

该研究基于7100万用户的手机记录和100万用户两年多来的位置跟踪信息，考察了中国广东省新冠肺炎大流行对劳动力市场的影响。中国广东省的GDP仅次于世界前12个国家。使用标准的差异中差异框架，我们的分析发现，从2020年1月底到2020年9月底，大流行使失业率增加了27-63%，远远高于官方统计数据。女性、40岁以上的工人和移民受影响更大。在那些更依赖出口、或酒店行业在GDP中所占比重较高、但在金融和医疗保健行业所占比重较低的城市，失业率的上升更为明显。在全面重新开放5个月后，就业尚未恢复到大流行前的水平。这种挥之不去的影响反映了这一大流行病的全球性质和世界经济的相互关联

**Abstract**

Based on mobile phone records for 71 million users and location tracking information for one million users over two years, this study examines the labor market impacts of the COVID-19 pandemic in China's Guangdong province, whose GDP is larger than all but the top 12 countries in the world. Using a standard difference-in-differences framework, our analysis finds that the pandemic increased unemployment by 27-63% from late January 2020 to the end of September 2020, much higher than the official statistics. Females, workers older than 40, and migrants are more affected. Cities that rely more on export or have a higher share of GDP in the hospitality industry but a lower share in the finance and healthcare industries experienced a more pronounced increase in unemployment. Five months after the full reopening, employment has not recovered to the pre-pandemic level. This lingering impact reflects the global nature of the pandemic and the interconnectedness of the world economy.

Covid-19 → 失业率

Keywords: COVID-19, Unemployment, Mobile phone

JEL Codes: E24, H12

* Corresponding author: Yongheng Deng: University of Wisconsin-Madison (4110 Grainger Hall, 975 University Ave, Madison, WI 53706, U.S.A., tel: +1 (608) 262-4865, e-mail: yongheng.deng@wisc.edu). Teng Li: Sun Yat-sen University, (e-mail: liteng27@mail.sysu.edu.cn); Panle Jia Barwick: Cornell University, (e-mail: panle.barwick@cornell.edu);; Xinfei Huang: Sun Yat-sen University, (e-mail: huangxf3@mail.sysu.edu.cn); Shanjun Li: Cornell University (e-mail: sl2448@cornell.edu). We thank Lihua Deng and Jiazhu Pan for generous help with data and Chris Chen for the literature review. P. Barwick and S. Li gratefully acknowledge the support by Cornell Center for Social Sciences and Cornell Atkinson Center for Sustainability. X. Huang gratefully acknowledges the support of the Major Project of the National Social Sciences Fund (16ZDA042). T. Li gratefully acknowledges the support of the Youth Program of the National Natural Science Foundation of China (72003202).

应对COVID-19大流行对经济的不利影响的有效和有针对性的政策有赖于对不同人口群体和地理区域的劳动力市场影响的迅速和准确衡量。劳动力市场结果的传统衡量标准，特别是失业率，都是基于调查。除了很大的时间滞后和小地理区域的可用性有限之外，从调查中推断出的统计数据也有很大的不确定性，经常要进行修订

在中国，失业信息来源于2018年之前在失业救济机构登记的人数，并辅以2018年之后的入户调查由于很大一部分人口没有当地户口，因此被排除在失业调查之外，因此准确测量失业率尤其具有挑战性。此外，报告和聚合错误以及潜在的数据操纵也被记录在案(贾尔斯等人，2005;刘,2012;蔡等人，2013)。在过去20年里，中国的全国失业率在3.1%-4.3%的狭窄范围内波动，这导致了对其可靠性的质疑(Feng, Hu, and Moffitt, 2017)，特别是在面对疫情带来的前所未有的社会和经济快速变化时

## 1. Introduction

Effective and targeted policies to address the adverse consequences of the COVID-19 pandemic for the economy rely on prompt and accurate measures of the labor market effects across different demographic groups and geographic regions. Traditional measures of labor market outcomes, in particular unemployment rates, are based on surveys. In addition to the substantial time lag and limited availability for small geographic areas, statistics inferred from surveys suffer from considerable uncertainty and are routinely revised.[1]

In China, information on unemployment is derived from the number of individuals who registered with the unemployment benefit agencies prior to 2018 and supplemented by household surveys afterward.[2] Measuring unemployment accurately is particularly challenging due to a large fraction of the population who do not have local household registrations (*Hukou*) and hence excluded from the unemployment surveys. Besides, reporting and aggregation errors, as well as potential data manipulations, have also been documented (Giles et al., 2005; Liu, 2012; Cai et al., 2013). China's national unemployment rate varies between a tight range of 3.1%-4.3% over the past two decades leading to questions about its reliability (Feng, Hu, and Moffitt, 2017), especially in the face of rapid and unprecedented social and economic changes brought about by the pandemic.[3]

This study leverages high-frequency and high-resolution mobile phone usage data in Guangdong, the most populous province in China, with a GDP larger than all but the top 12 countries in the world. Our primary data source consists of mobile phone records for 71 million users and location tracking information for one million randomly selected users from January 2018 to September 2020. We examine the pandemic's labor market impacts for various demographic groups and across cities with different industrial structures

[1] In addition to the delay, it is technically challenging to measure unemployment accurately. Their quality also varies considerably over time due to changes in participation rates, modifications in the survey methodology, inconsistencies and measurement errors in sample responses, or rotation group bias (Poterba and Summers, 1986; Jones and Riddell, 1999; Card, 2011; Feng and Hu, 2013; Meyer et al., 2015; Krueger et al., 2017; Heffetz and Reeves, 2019.).

[2] Please refer to this report (https://www.marketwatch.com/story/china-unveils-urban-survey-unemployment-rate-2018-04-17-1485030).

[3] Please also refer to an acticle in the Wall Street Journal (https://www.wsj.com/articles/chinas-jobs-rebound-doesnt-appear-as-robust-as-the-government-claims-11591551390).

这项研究利用了广东的高频和高分辨率手机使用数据。广东是中国人口最多的省份，其GDP超过世界上除前12个国家外的所有国家。我们的主要数据来源包括7100万用户的手机记录和2018年1月至2020年9月随机选择的100万用户的位置跟踪信息。我们采用标准差异中的差异(DID)框架，研究了大流行对不同人口群体和不同产业结构城市的劳动力市场影响。我们利用两个独特的数据特征来确定对失业的影响:a)通过电话联系指定失业救济机构的人数;b)在现场工作的人数。

我们的分析得出了几个关键发现。首先，根据以下两种衡量失业影响的策略，截至2020年9月30日，疫情已使广东的失业率上升了27-63%。失业率的急剧上升远远高于政府统计数据所显示的广东省同期13.3%的失业率(从1 - 3月的2.26个百分点上升到7 - 9月的2.56个百分点)。

二是广东全面复工近5个月来，到失业救济机构申请失业救济的人数持续上升，就业没有恢复到疫情前水平。疫情对失业的冲击在不同人口群体中极不均匀，在女性、40岁以上的人、特别是移徙者中尤为明显。在我们的抽样期间，移徙者失业人数的不断增加没有显示出减少的迹象。这与国家统计局报告的农民工数量大幅减少相呼应，表明这一群体有可能大规模裁员。

by employing the standard difference-in-differences (DID) framework. We leverage two unique data features to bound the impact on unemployment: a) the number of individuals who contact the designated unemployment benefits agencies via phone, and b) the number of individuals working on-site.[4]

Several key findings emerge from our analysis. First, the pandemic has increased unemployment in Guangdong by 27-63% as of September 30, 2020, based on two strategies to measure the unemployment impact described in detail below. The sharp rise in unemployment is much higher than the government statistics that report an increase of 13.3% in Guangdong province's unemployment rate (from 2.26 percentage points in January-March to 2.56 percentage points in July-September) during the same period.

Second, nearly five months after Guangdong's full reopening, the number of individuals who reach out to unemployment benefits agencies to claim unemployment benefits continues to rise, and employment has not returned to the pre-pandemic level. The pandemic's shock on unemployment is highly uneven across demographic groups and more pronounced among females, people over 40, and especially migrants. The escalating increase in unemployment among migrants shows no sign of abatement during our sample period. This echoes a massive reduction in NBS' reported migrant workers and indicates the possibility of a large-scale layoff among this group.

Third, the pandemic's impact is more substantial in cities with a high labor share of hospitality, real estate, or transportation industries but less severe in cities where employments are concentrated in finance, health care, or education industries. In addition, the impact is more pronounced in cities that rely heavily on export, reflecting the global nature of the shock in an interconnected world economy. Industry compositions account for 36% of the heterogeneity in the pandemic's unemployment impact across cities, while trade exposure contributes to 30% of the heterogeneity.

第三，在酒店、房地产或交通运输行业劳动力占比高的城市，疫情的影响更大，但在就业集中于金融、医疗保健或教育行业的城市，疫情的影响不那么严重。此外，严重依赖出口的城市受到的影响更为明显，反映出在相互关联的世界经济中，此次冲击具有全球性。在大流行跨城市失业影响的异质性中，行业构成占36%，而贸易敞口占30%。

[4] Researchers in recent studies have used mobile phone data to improve labor market measurements (Toole et al., 2015; Barwick et al., 2019), track human movement in real-time and at a fine spatial scale, and quantify the pandemic's impact on voting behaviors, mobility, and social contacts (Gonzalez et al., 2008; Ahas et al., 2010; Couronné et al., 2013; Chen et al., 2018; Kreindler and Miyauchi, 2019; Chen et al., 2019; Atkin et al., 2020; Chen and Pope, 2020; Chevalier et al., 2020; Gupta et al., 2020).

最后，我们的研究结果揭示了大流行对劳动力市场影响的严重性和不均匀性，这说明了在细粒度层面进行分析的重要性。此外，这些结果表明，疫情通过供应链和贸易渠道在一个国家的城市和世界各国之间产生了涟漪效应。一个城市(或一个国家)的产业构成、贸易敞口和供应链性质是决定大流行对其经济影响的关键因素。我们的措施帮助我们从微观层面了解大流行对劳动力市场的影响，并为有针对性的政策提供信息，以帮助受影响最严重的群体和地区。

Lastly, our results unmask the severity and uneven impact of the pandemic's labor market implications, which speak to the importance of conducting analysis at granular levels. In addition, these results illustrate the rippling effect of the pandemic across cities within a country and countries worldwide through the supply chain and the trade channels. The industry composition of a city (or a country), exposure to trade, and the nature of the supply chain are crucial factors in determining the pandemic's effect on its economy. Our measures help us understand the pandemic's labor market impact at a granular level and inform targeted policies to help the most severely affected groups and regions.

## 2. Background and Data

### *2.1 Background and Data sources*

Exploiting the increasingly available high-frequency and high-resolution mobile phone data is particularly advantageous for China, as its cellphone penetration rate is high among developing countries. According to the 2018 China Family Panel Studies, a nationally representative longitudinal survey of individuals' social and economic status, 89% of correspondents sixteen years and older reported possessing a cellphone. In addition, each household owns 2.5 cell phones on average (National Bureau of Statistics 2018). Appendix Figure A.1 shows a strong correlation between the number of China Mobile users and the number of residents by city. Cities with a higher GDP per capita (represented by the size of the circles in Figure A.1) tend to have higher mobile phone ownership.

The context of our analysis is Guangdong, the most populous province with the largest provincial GDP in China. Guangdong contributes to 11% of China's GDP and around a quarter of China's foreign trade (China Statistical Yearbook 2020). Its major cities include Shenzhen and Guangzhou, among the wealthiest and economically most advanced cities in China. Its economy is widely recognized as the most dynamic and resilient among all provinces in China (Yu and Zhang, 2011; Gong et al., 2020). Another reason that makes Guangdong relevant is that the number of daily confirmed COVID cases are under a few handfuls since the full reopening (Appendix Figure A.2), similar to most other provinces in China. Our measures on the pandemic's consequences could apply to other regions as well.

Figure A.1: Mobile users vs. residents in 2019
Average daily mobile users in 2019 (million)
Residents in 2019 (million)
0
5
10
15
20
0
5
10
15
20
Yu
Gu
Ch
Ya
Sh
Ma
Ji
Ha
Ch
Ch
Ch

Figure A.2. Newly confirmed COVID-19 cases in Guangdong province

我们的数据来自中国移动，中国主要的手机服务提供商。我们获得了2018年1月至2020年9月广东省7100万用户的详细手机使用情况(加密的主叫和接收方id、通话日期和通话时长，以秒为单位)，占全省移动用户总数的63%。我们观察一些用户人口统计信息，例如年龄、性别和电话号码的注册地点。此外，我们每隔5分钟就可以访问同一时间段内随机选择的100万用户的位置记录。

Our data come from China Mobile, the dominant cellular service provider in China. We have access to detailed phone usage (encrypted IDs of the calling party and the receiving party, date of calls, and call duration in seconds) for all of its 71 million users in Guangdong Province from January 2018 to September 2020, accounting for 63% of all mobile users in the province. We observe some user demographic information, such as age, gender, and the place where the phone number is registered. In addition, we have access to the location records every five-minute interval for one million randomly selected users during the same period.

Figure A.3. Correlation of Baidu index between keywords of "12333" and "unemployment insurance" in Guangdong province

我们利用手机数据的两个特征来构建失业测度。第一个特点是详细的呼叫记录(每次呼叫的时间和时长)，拨打指定的政府失业救济热线(12333)。该热线提供全面的一站式社会保险公共服务，提供资格信息，帮助失业登记，方便申请失业救济。相对于网上申报或到当地的社会保障局，拨打12333热线是许多人的首选，因为它简单，客服提供全方位的帮助。图A.3为广东省2019 - 2020年关键字“12333”和“失业保险”的周百度指数。6两个关键字百度指数在样本周期内的相关性为0.83。这两个关键字的索引的联用为使用12333热线作为申请失业救济的个人的代理提供了额外的支持。

### *2.2 Unemployment measures*

We leverage two features of the mobile phone data to construct unemployment measures. The first feature is detailed call records (time and duration of each call) to the designated government hotline (12333) for unemployment benefits. The hotline offers comprehensive one-stop social insurance public service, provides eligibility information, helps with unemployment registration, and facilitates applications for unemployment benefits. Relative to filing online or visiting local social security bureaus, calling the designated hotline 12333 is the preferred choice for many due to its simplicity and all-inclusive help from customer services.[5] Appendix Figure A.3 shows the weekly Baidu Index for the keywords – “12333” and “unemployment insurance” – in Guangdong province from 2019 to 2020.[6] The correlation of the Baidu Index of the two keywords is 0.83 during the sample period. The co-movement of the index for the two keywords offers additional support for using the 12333 hotline as a proxy for individuals claiming unemployment benefits.

尽管这条热线很受欢迎，但拨打12333的人数为失业水平提供了一个较低的估计值，原因与政府官方统计数据保守的原因相同。并不是所有的失业人员都向政府机构申请失业救济，尤其是那些对很快找到新工作持乐观态度的人。此外，中国的终身失业救济金的上限是24个月，因此限制了过去已经用光救济金的人的选择。因此，我们下面的分析不是关注失业电话的水平，而是利用它的变化。我们表明，对失业电话变化的度量可以提供短期劳动力市场动态的有用信息，否则无法通过官方统计数据获得。

Despite the popularity of the hotline, the number of individuals making calls to 12333 provides a *lower bound* estimate for the *level* of unemployment, for the same reasons that make the government official statistics conservative. Not all unemployed workers reach out to government agency to claim unemployment benefits, especially those who are optimistic about finding a new job soon. In addition, the

[5] According to a 2016 study on China’s special economic zones, 96% of individuals in Xiamen city use 12333 to claim unemployment office, and only 4% visit local social security bureaus or file online (Tao and Yuan, 2016).

[6] Baidu Index, similar to Google Trends, is a keyword-analysis tool launched by Baidu, the largest search engine company in China. It reflects the search frequency of specific keywords on the Baidu website.

Figure A.4. Unemployment rate by city in 2019 based on the number of individuals making unemployment calls

由于人们可能多次拨打热线电话申请失业救济，我们使用拨打12333的人数，而不是拨打12333的人数，来构建我们的第一个失业指标。为了简单起见，术语“拨打12333的个人数量”和“失业电话数量”在整个分析中交替使用。图A.4绘制了2019年跨城市拨打12333的人数。城市一级的失业电话与国家统计局发布的官方失业率之间的相关性相当高，2019年为0.7。国家统计局每年只发布城市一级的统计数据。

lifetime unemployment benefits in China are capped at 24 months, thus limiting choices for people who have exhausted benefits in the past. Therefore, instead of focusing on the level of unemployment calls, our analysis below exploits its *changes*. We show that the measure of changes in unemployment calls can provide useful information on short-run labor market dynamics otherwise unavailable through official statistics.

As people might reach out to the hotline multiple times to claim their unemployment benefits, we use the number of individuals calling 12333, instead of the number of calls to 12333, to construct our first unemployment measure. For brevity, the term "number of individuals calling 12333" and "number of unemployment calls" are used interchangeably throughout the analysis. Appendix Figure A.4 plots the number of individuals calling 12333 across cities in 2019. The correlation between city-level unemployment calls and the official unemployment rate released by the NBS, which is only available annually for city-level statistics, is reasonably high at 0.7 in 2019.

The second feature of the mobile phone data is the location tracking information (in longitude and latitude) collected by mobile devices during 5-minute intervals except when they are powered off.[7] We randomly select one million mobile users and use their location information at 5-minute intervals from January 2018 to September 2020 to construct their job and home locations. We define the work location as the location where a user spends at least 5 hours a day between 9 am and 6 pm for at least fifteen workdays in a given month. The home location is similarly constructed, except that we use the location with the longest duration between 10 pm and 7 am each month.[8] These geocoded locations trace out individuals' spatial trajectories over time and allow us to record the time of arrival and departure at job locations.

We use reductions in the number of people working on-site before and after the lockdown and relative to 2019 as our second measure of pandemic-induced unemployment. Changes in commuting patterns,

[7] Researchers have used such mobile positioning data to study urban and transportation issues (Gonzalez et al., 2008; Ahas et al., 2010), though few studies exploited long panels of location data to examine labor market dynamics (Barwick et al., 2019).

[8] Location information from 7 am - 9 am and 6 pm -10 pm is discarded because people are likely on the move during these time intervals.

手机数据的第二个特点是，除关机外，移动设备每隔5分钟收集的位置追踪信息(经度和纬度)。我们随机选择100万移动用户，在2018年1月至2020年9月期间，每隔5分钟使用他们的位置信息构建他们的工作和家庭位置。我们将工作地点定义为用户在给定的月中至少15个工作日中每天从上午9点到下午6点花费至少5个小时的位置。家庭位置的构造与此类似，除了我们使用的位置在每个月的晚上10点到早上7点之间的最长时间。这些地理编码的位置可以追踪个人随时间变化的空间轨迹，让我们可以记录到达和离开工作地点的时间。

我们将封锁前后现场工作人数的减少以及与2019年相比的减少作为大流行导致失业的第二个衡量标准。通勤方式的变化，尤其是持续的变化，为失业变化提供了另一个有价值的晴雨表，特别是在失业福利计划的参与率较低的情况下(就像中国的情况一样)。鉴于其中一些变化反映了封锁后更灵活的工作模式，如远程办公，因此应将其解释为对大流行导致的失业的上限估计。然而，我们在下面提供的证据表明，我们基于长期通勤模式对失业的衡量不太可能是由远程办公驱动的。附录包含了更多关于我们的样本大小的细节，以及我们如何构建各种衡量劳动力市场结果的方法。

especially on a continuing basis, provides another valuable barometer of changes in unemployment, especially when participation in the unemployment benefit programs is low (as is the case in China). To the extent that some of these changes reflect more flexible work modes post the lockdown, such as telecommuting, they should be interpreted as an upper bound estimate on pandemic-induced unemployment. However, we provide evidence below that our measure of unemployment based on commuting patterns over an extended period of time is unlikely to be driven by telecommuting. The appendix contains more details on our sample size and how we construct various measures for the labor market outcome.

### *2.3 Sample construction*

在我们使用失业电话的分析中，我们使用拨打12333热线的人数，而不是拨打12333的人数，以解决人们可能多次拨打热线申请失业救济金的问题。换句话说，我们只计算用户第一次拨打失业救济热线的次数。当我们测量到热线的呼叫时长时，我们会聚合所有后续呼叫的持续时间。我们的主要分析排除了18岁以下的用户，因为18岁以下的个人不太可能在《未成年人保护法》的约束下工作。排除25岁以下的用户(为了排除仍在上学的用户)的结果几乎相同(见附录)。

In our analysis using unemployment calls, we use the number of individuals making calls to the 12333 hotline, instead of the number of calls to 12333, to address the issue that people might reach out to the hotline multiple times to claim their unemployment benefits. In other words, we only count the first time when a user reaches out to the unemployment benefit hotline. We aggregate the duration of all subsequent calls when we measure call duration to the hotline. Our main analysis excludes users under the age of 18, as the individuals under the age of 18 are unlikely to be working subject to the Law on Protection of Minors. Results excluding users under the age of 25 (to eliminate those still in school) are almost identical (see appendix).

我们的一些分析考察了移民和非移民之间的异质劳动力市场前景。值得注意的是，从2014年起，一直在广东工作但没有广东户口的外来务工人员可以享受失业救济。这在很大程度上是由广东省政府设计的，目的是吸引外来人口，并帮助改善劳资关系。由于我们不观察个人是否有当地户口(官方对流动人口的定义)，我们将流动人口定义为在广东省以外注册电话号码的个人。这是对当地户口状况的不完美衡量。在我们的分析中，广东以外的工人可以在广东购买和注册他们的手机，并被视为非外来人口。因此，农民工和居民之间的实际失业差距可能比我们估计的还要大。

Some of our analyses examine heterogeneous labor market prospects between migrants and non-migrants. It is important to note that migrants who had been working in Guangdong but did not have Guangdong *Hukou* became eligible for unemployment benefits since 2014.[9] This was largely designed by the Guangdong government to attract migrants and to help improve labor relations. As we do not observe whether an individual has local *Hukou* status – the official definition of migrants, we define migrants as individuals who registered their phone numbers outside Guangdong province. This is an imperfect measure of local *Hukou* status. Workers from outside Guangdong can buy and register their mobile phones in

[9] Please refer to the announcement by Human Resources and Social Security Department in Guangdong, https://www.gdhrss.gov.cn/sy/20140801/10101.html.

Guangdong and be treated as non-migrants in our analysis. Consequently, the actual gap in unemployment between migrants and residents might be even larger than our estimates.

广东封锁从2020年1月23日至2月24日，持续了32天。2020年2月24日，广东进入第一期复工复产，复工复产持续76天。在第一阶段重新开放期间，人们被允许(在某些行业，人们被鼓励)返回工作岗位，参观户外公共场所。第二阶段重新开放，即全面重新开放，于2020年5月9日正式开始。包括购物中心、超市和餐馆在内的所有企业都被允许全面重新开业。唯一的例外是一直关闭到7月中旬的电影院。我们将封锁前60天到封锁后252天的时间间隔划分为四个时期:封锁前(60天)、封锁期间(32天)、第一阶段重新开放(76天)和第二阶段完全重新开放(144天)。

### *2.4 Descriptive analysis*

The lockdown in Guangdong lasted 32 days, from January 23 to February 24, 2020. On February 24, 2020, Guangdong province entered Phase I reopening, which lasted 76 days. During Phase I reopening, people were allowed (and in certain industries, people were encouraged) to go back to work and visit outdoor public places. Phase II reopening, or full reopening, officially started on May 9, 2020. All businesses, including shopping malls, supermarkets, and restaurants, were allowed to reopen fully. The only exception was movie theaters that remained closed till mid-July. We delineate the interval from the 60 days before the lockdown to the 252 days after the lockdown into four periods: before lockdown (60 days), during the lockdown (32 days), Phase I reopening (76 days), and Phase II full reopening (144 days).

Our analyses use a standard DID framework and exploit differences in our two key measures of labor market outcomes between 2020 (the treatment group) and 2019 (the control group). The event date for 2020 is January 23, 2020, while the event date for 2019 is February 3, 2019, two days before the Chinese New Year in 2019.[10] In other words, our analyses compare changes in labor market outcomes before and after the event date in 2020 with changes in labor market measures before and after the exact event date in 2019.

Figure 1 depicts the differences in the daily number of individuals calling 12333 between 2019 and 2020. Red bars indicate more people calling 12333 in 2020 than in 2019, and blue bars indicate fewer people calling 12333 in 2020 than in 2019. The daily number of individuals calling 12333 in 2020 was similar to the 2019 baseline before the lockdown, which supports the parallel trend assumption underlying the DID strategy. During the lockdown, it dropped significantly due to the uncertainty about the severity and duration of the epidemic during the pandemic's initial stage. As the severity of the pandemic unfolded in China, the number of individuals calling 12333 increased sharply one month after the beginning of Phase

我们的分析使用了一个标准的DID框架，并利用了2020年(治疗组)和2019年(对照组)之间劳动力市场结果的两个关键指标的差异。2020年的活动日期是2020年1月23日，而2019年的活动日期是2019年2月3日，也就是2019年中国新年的前两天。换句话说，我们的分析比较了2020年事件发生前后劳动力市场结果的变化与2019年确切事件发生前后劳动力市场指标的变化。

[10] As Guangdong's lockdown occurred on January 23, two days before the Chinese New Year on January 25, the most important holiday in China, it is critical to control this holiday effect directly. Hence, we use the lunar calendar instead of the standard almanac calendar to define the event date.

图1描述了2019年和2020年之间每天拨打12333的人数的差异。红色条表示2020年拨打12333的人数比2019年多，蓝色条表示2020年拨打12333的人数比2019年少。2020年每天拨打12333的人数与封锁前的2019年基线相似，这支持了作为DID战略基础的平行趋势假设。在封锁期间，由于大流行初期对疫情的严重程度和持续时间的不确定性，该指数大幅下降。随着疫情在中国的严重程度，在第一阶段重新开放开始一个月后，拨打12333的人数急剧增加，这一趋势一直持续到我们的数据周期结束。17黑线虚线表示2019年至2020年拨打12333的人数累积差异。疫情前60天的累计差异徘徊在零附近，表明大流行前两年的劳动力市场前景具有可比性。

I reopening, and the trend continued to the end of our data period.[11] The dotted black line shows the cumulative differences in individuals calling 12333 between 2019 and 2020. The cumulative differences during the first 60 days prior to the event hover around zero, suggesting that the labor market prospects are comparable across the two years before the pandemic.

从数量上看，广东省每天拨打12333的平均人数为11828人，四分位数区间为11371人。从封锁到2020年9月30日，总共有331万人拨打12333，而2019年同期的数字是258万人。这些数字表明，在我们的样本期间，大流行已导致失业人数累计增加28%。

In terms of magnitude, the average number of individuals calling 12333 per day in Guangdong province is 11,828, with an interquartile range of 11,371. There are a total of 3.31 million individuals calling 12333 from the lockdown to September 30, 2020, while the number during the same period in 2019 is 2.58 million. These numbers suggest that the pandemic has led to a cumulative 28% increase in unemployment in our sample period.

我们的第二套劳动力市场指标包括通勤上班的人数，以及在100万随机选择的移动用户中有条件地在工作场所工作的时间。2019年日均通勤人数为48万人，与广东省劳动人口比例(《广东省统计年鉴2020》)一致。图2展示了2019年和2020年的每日差异。与图1类似，2019年和2020年之间的通勤模式在封锁前几乎没有变化，无论是在每日通勤人数还是在现场工作时间。大流行前，每天通勤的平均人数为484,750人(四分位数区间为124,690人)，平均工作时数为8.8小时(四分位数区间为0.6小时)。

Our second set of labor market measures includes the number of people commuting to work and the amount of time spent at the workplace conditional on working on-site among one million randomly chosen mobile users. The average number of daily commuters is 0.48 million in 2019, consistent with the working population ratio in Guangdong (Guangdong Provincial Statistical Yearbook 2020). Figure 2 presents the daily differences between 2019 and 2020. Similar to Figure 1, the commuting patterns between 2019 and 2020 displayed few differences before the lockdown, both in terms of daily commuters and hours worked on-site. The average number of daily commuters is 484,750 (with an interquartile range of 124,690) and the average hours worked are 8.8 (with an interquartile range of 0.6) pre the pandemic.

During the lockdown in 2020, both the number of commuters and hours worked on-site registered a sharp reduction relative to those during the same lunar calendar days in 2019, as expected. The drops reached their trough toward the end of the lockdown, with the number of commuters decreasing by 0.19 million (about 40%) and the number of hours worked on-site falling by two hours per day. Both series had gradually improved since Phase I reopening. By September 2020, 144 days after Phase II full reopening, the number of hours working on-site had returned to the 2019 level. However, the number of commuters

[11] Some of the time-series variations in the bar chart reflect differences between weekdays and weekends and during holidays. For example, few people reached out to the unemployment hotline right before the Phase II reopening, which coincides with International Labor Day, a 5-day public holiday from May 1 to May 5 (event day 100 to event day 104).

正如预期的那样，在2020年封锁期间，通勤人数和现场工作时间都比2019年同期大幅减少。在封城结束时，下降达到了最低点，通勤人数减少了19万(约40%)，每天现场工作时间减少了2小时。自第一期重新开放以来，这两个系列的情况已逐步改善。到2020年9月，在第二阶段全面重新开放144天之后，现场工作小时数已恢复到2019年水平。然而，通勤人数还没有完全恢复到大流行前的水平。2020年和2019年之间的差距一直持续到样本结束，2020年9月通勤上班的人数比2019年9月减少了6.7%。

has not fully recovered to the pre-pandemic level. The gap between 2020 and 2019 persists till the end of our sample, with 6.7% fewer people commuting to work in September 2020 relative to September 2019.

## 3. Empirical framework

我们的分析采用了标准的差中差(DID)方法，将2020年疫情发生前后(广东实施封城措施时)的劳动力市场结果与2019.12年相同(农历)日历日期前后的劳动力市场结果进行了比较

Our analysis employs the standard difference-in-differences (DID) approach by comparing labor market outcomes in 2020 before and after the event date (when Guangdong implemented the lockdown) with those before and after the same (lunar) calendar dates in 2019.[12]

以2020年为治疗组，2019年为对照组。为了控制时间变化的不可观察的潜在差异，我们包括一组丰富的固定效应，如工作日、事件日、假日和治疗组固定效应。识别假设是，在包括这些对照之后，在没有疫情的情况下，两组之间不存在时间变化的不可观察的系统性差异。事件研究的结果支持两组在事件日期之前的共同趋势假设。我们使用以下DID框架和十天间隔追踪大流行随时间的动态影响:

We use the year 2020 as the treatment group and the year 2019 as the control group. To control for potential differences in time-varying unobservables, we include a rich set of fixed effects such as day-of-week, event-day, holiday, and the treatment group fixed effects. The identification assumption is that after including these controls, there are no systemic differences in time-varying unobservables between the two groups in the absence of the pandemic. Results from event studies support this common trend assumption between the two groups prior to the event date. We use the following DID framework and ten-day intervals to trace out the dynamic impact of the pandemic over time:

$$y_{cit} = \sum_{q=-5}^{24} \beta_q \cdot d_i \cdot \mathbf{1}(t \in [q*10+1, (q+1)*10]) + \alpha_c + \gamma_i + \eta_t + \xi_{it} + \varepsilon_{cit}, \qquad (1)$$

其中$cc$表示一个社区(蜂窝基站区域)，$ii$表示治疗组(2020年)或对照组(2019年)，$tt$表示事件日(t=0表示2020年1月23日和2019年2月3日)。事件窗口为封锁前60天和封锁后252天。

where $c$ denotes a neighborhood (a cell-tower area), $i$ denotes the treatment group (year 2020) or the control group (year 2019), and $t$ denotes the event-day (t=0 stands for January 23 in 2020 and February 3 in 2019). The event window is sixty days before the lockdown and 252 days post the lockdown.

是结果变量，例如调用12333的个人的日志数。我们基于log(结果+1)报告结果，以避免取对数除以零。然而，基于反双曲正弦函数(与对数函数非常相似，可以处理零值)的结果非常相似。$\beta\beta qq$是事件研究系数，捕捉治疗组和对照组之间的差异

$y_{cit}$ is the outcome variable, such as the log number of individuals calling 12333. We report results based on log(outcome+1) to avoid taking the logarithm over zero. However, results based on the inverse hyperbolic sine function (which is very similar to the log function and can handle zero values) are very similar. $\beta_q$ are event-study coefficients, capturing differences between the treatment group and the control group. Variable $d_i$ is a dummy, that is equal to one for the treatment group. $\mathbf{1}(.)$ is an indicator variable for

[12] As Guangdong's lockdown occurred two days before the 2020 Chinese New Year, we use the lunar calendar instead of the standard almanac calendar to define the event date. Specifically, the event date for 2020 is January 23, 2020, while the event date for 2019 is February 3, 2019, two days before the Chinese New Year in the lunar calendar.

组。变量$ddcc$是一个假变量，它等于治疗组的1。$\mathbf{1}(.)$为样本每10天间隔的指标变量。我们控制了邻里固定效果$\alpha\alpha cc$，群体固定效果$\gamma\gamma cc$和312事件日固定效果$\eta\eta cc$。我们还控制了假日固定效应$\xi\xi cccc$，它随群体和时间的不同而不同(例如，2019年和2020年的国际劳动节假期在不同的阴历日期)，以及每周的固定效应。标准错误集中在事件日级别。

为了进一步探索城市间的异质性以及产业构成和贸易敞口的重要性，我们采用了以下规范

each 10-day interval of the sample. We control for neighborhood fixed effects $\alpha_c$, group fixed effects $\gamma_i$, and 312 event-day fixed effects $\eta_t$. We also control for the holiday fixed effect $\xi_{it}$ , that varies by group and time (e.g., the International Labor Day holiday falls on different lunar calendar days in 2019 and 2020) as well as day-of-week fixed effects. Standard errors are clustered at the event-day level.

To further explore the heterogeneity across cities and the importance of industrial composition and trade exposure, we employ the following specification:

$$y_{cit} = d_i \cdot \mathbf{1}(t \in [0, 252]) \cdot \boldsymbol{Z}'\boldsymbol{\tau} + d_i \cdot \boldsymbol{Z}'\boldsymbol{\mu} + \mathbf{1}(t \in [0, 252]) \cdot \boldsymbol{Z}'\boldsymbol{\rho} +$$

$$\beta \cdot d_i \cdot \mathbf{1}(t \in [0, 252]) + \alpha_c + \gamma_i + \eta_t + \xi_{it} + \varepsilon_{cit}, \tag{2}$$

where $\boldsymbol{Z}$ is a vector of city attributes in 2019 and $\boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\rho}$ are corresponding coefficients. For example, $\boldsymbol{Z}$ could be a city's labor share in each of the 13 major industries, dummies for the 21 cities, or a city's export-over-GDP ratio. In addition to the interaction between the pandemic treatment and city attributes, we control for all lower-level interactions in the regression. Variables $d_i$, $\mathbf{1}(.)$ and the fixed effects $\alpha_c, \gamma_i, \eta_t, \xi_{it}$ are the same as in equation (1). The key coefficient is $\boldsymbol{\tau}$, which measures the heterogeneous impact by city characteristics $\boldsymbol{Z}$ based on their values in 2019. Unlike in equation (1), where we estimate the pandemic's impact for each ten-day interval, here we estimate the average effect $\boldsymbol{\tau}$ over all periods and focus on heterogeneity across industries and cities.

其中$ZZ$为2019年城市属性向量，$\tau\tau$、$\mu\mu$、$\rho\rho$为对应系数。例如，$ZZ$可以是一个城市在13个主要产业中的劳动占比，21个城市的假数字，或者一个城市的出口占gdp的比例。除了大流行治疗和城市属性之间的相互作用外，我们还控制回归中所有较低级别的相互作用。变量$ddcc$，$\mathbf{1}(.)$固定效应$\alpha\alpha cc$、$\gamma\gamma cc$、$\eta\eta cc$、$\xi\xi cccc$与式(1)相同，关键系数为$\tau\tau$，根据其2019年的数值衡量城市特征$ZZ$的异质性影响。在公式(1)中，我们估计了大流行每10天间隔的影响，而在这里，我们估计了所有时期的平均影响$\tau\tau$，并关注行业和城市之间的异质性。

The regression analyses using the number of commuters are analogous, except that the observation is at the neighborhood and time-window level. For example, for the commuter definition that uses two weeks as the relevant time window, the observation is at the neighborhood-fortnight window. We drop the day-of-week fixed effects, replace event-day fixed effects with event-fortnight fixed effects and cluster the standard errors at the fortnight level.

使用通勤者数量的回归分析是类似的，除了观察是在邻居和时间窗水平。例如，对于使用两周作为相关时间窗口的通勤者定义，观察是在社区-两周窗口。我们去掉了一周的固定效果，用事件两周的固定效果代替了事件一天的固定效果，并将标准误差聚类到两周级别。

## 4. Main results

Figure 1 and Figure 2 provide compelling evidence that the pandemic has greatly affected the labor market in 2020. We now use equation (1) to quantify the pandemic's effect on unemployment. All regressions

图1和图2提供了令人信服的证据，表明疫情对2020年的劳动力市场产生了巨大影响。我们现在使用方程(1)来量化大流行对失业的影响。所有回归均包括事件日、工作日、假日、治疗组和社区固定效应。与上面关于数据模式的讨论类似，我们的回归分析也包括两部分。第一部分利用了打失业电话的人数的变化(或简单地说“失业电话”)，而第二部分利用了通勤模式的变化。

include event-day, day-of-week, holiday, treatment group, and neighborhood fixed effects. Similar to the discussion on data patterns above, our regression analysis also consists of two parts. The first part exploits variation in the number of individuals making unemployment calls (or “unemployment calls” for brevity), while the second part leverages changes in commuting patterns.

### *4.1 Effect on unemployment calls*

为了检验失业电话的变化，我们总共使用了489,514个邻里日观察数据。图3的面板(a)描述了2020年拨打12333的人数相对于2019年的人数变化百分比。事件前系数估计集中在零附近，在统计上不显著，支持我们对2019年和2020年之间平行趋势的识别假设。与图1中的描述性证据一致，失业电话在封锁期间下降，但在第一阶段重新开放一个月后急剧增加。这一趋势一直持续到9月底，即第二阶段重新开放近5个月后。从3月中旬到9月，失业求助电话的增长幅度在20-50%之间。

For regressions that examine changes in unemployment calls, we utilize a total of 489,514 neighborhood-day observations. Panel (a) of Figure 3 depicts the percentage change in individuals calling 12333 in 2020 relative to those in 2019. The pre-event coefficient estimates center around zero and are statistically insignificant, supporting our identification assumption of the parallel trend between 2019 and 2020. Consistent with the descriptive evidence in Figure 1, unemployment calls dropped during the lockdown but increased sharply one month after the Phase I reopening. The trend continued till the end of September, nearly five months into the Phase II reopening. The increase in the number of unemployment calls ranges from 20-50% from mid-March to September.

Panel A of Table 1 presents the coefficient estimates of $\beta_q$ in equation (1), except that the ten-day intervals are grouped into four periods: before lockdown, during the lockdown, Phase I reopening, and Phase II full reopening. Column (1) shows the pandemic’s impact on the number of unemployment calls, while column (2) examines call duration. Echoing results in Figure 3, the number of individuals calling 12333 decreased by 37% during the lockdown and increased by 25% during Phase I reopening and 45% during Phase II reopening. The call duration displayed a similar pattern: the average call time dropped during the lockdown but increased after the reopening. Call duration increases partly because more migrants applied for unemployment benefits after the reopening (as we show below in heterogeneity analysis) and that they generally need to provide more information than local residents.[13]

Panel B of Table 1 further groups the three periods during and post the lockdown into one group. The coefficient estimates directly measure the pandemic’s average labor market effect. Overall, the pandemic

[13] The average duration per call for migrant and non-migrant workers are about 6 and 4 minutes, respectively.

表1的面板A给出了方程(1)中$\beta\beta qq$的系数估计值，但10天的间隔被分为四个时期:封锁前、封锁期间、第一阶段重新开放和第二阶段全面开放。第(1)栏显示了疫情对失业电话数量的影响，而第(2)栏审查了电话通话时长。与图3的结果相呼应的是，在封锁期间，拨打12333的人数减少了37%，在第一阶段重新开放时增加了25%，在第二阶段重新开放时增加了45%。通话时长显示出类似的模式:在封锁期间平均通话时间下降，但在重新开放后增加。通话时长增加的部分原因是，在重新开放后，更多的移民申请了失业救济(如下面的异质性分析所示)，他们通常需要提供比当地居民更多的信息

表1的面板B将锁定期间和锁定后的三个时间段进一步分组。系数估计直接衡量了大流行对劳动力市场的平均影响。总体而言，疫情使失业电话增加了27%。这与我们在比较2020年至2019年间拨打12333的原始累积人数时所讨论的幅度非常相似，这表明混淆因素(如我们丰富的固定效果和控制所捕捉到的)的作用是有限的。如上所述，失业电话数量是失业人数的下限估计，因为不是所有失业的人都申请失业救济。然而，鉴于2020年和2019年在大流行之前的失业求助电话数量非常相似，表1中估计的失业求助电话变化百分比(27%)可能是大流行造成的失业率变化百分比的可靠衡量标准。

has increased the number of unemployment calls by 27%. This is very similar to the magnitude discussed above when we compare the raw cumulative number of individuals calling 12333 between 2020 and 2019, suggesting that the role of confounders (as captured by our rich set of fixed effects and controls) is limited. As discussed above, the level of unemployment calls is a lower bound estimate of the number of unemployed, as not all individuals who have lost jobs file for unemployment benefits. However, in light of the remarkable similarity in unemployment calls between 2020 and 2019 prior to the pandemic, the percentage *change* in unemployment calls estimated in Table 1 (27%) is likely a reliable measure of the percentage *change* in the unemployment rate as a result of the pandemic.

### *4.2 Heterogeneity analysis*

为了检验大流行在不同人口群体之间的不同影响，我们重复了按性别、年龄和移徙者身份和图3(b)-(d)组图系数进行的事件研究分析。具体而言，因变量是女性和男性(分组b)、40岁及以上的人和40岁以下的人(分组c)以及移徙者和非移徙者(分组d)之间的对数结果变量的差异。女性受大流行病的影响更大。从4月至9月，女性的失业电话数量比男性高出10-15个百分点。14由于许多托儿设施在疫情期间(部分)关闭，对家庭托儿服务的需求增加可能是女性失业影响更为明显的原因之一。

To examine the pandemic's differential impacts across demographic groups, we repeat the event study analysis by gender, age, and migrant status and plot coefficients in panels (b)-(d) of Figure 3. Specifically, the dependent variable is the difference in the logarithm outcome variable between females and males (panel b), between individuals 40 years old and above and those under 40 (panel c), and between migrants and non-migrants (panel d). Females are more affected by the pandemic. The number of unemployment calls by females is 10-15 percentage points higher than that by males from April to September.[14] As many childcare facilities were closed during (part of) the pandemic, the increasing demand for home childcare might have played a role in the more pronounced unemployment impact among females.

Older workers fared worse than younger cohorts: the number of unemployment calls by workers 40 and above is around 20 percentage points higher than those under 40. Lastly, migrants are most severely affected by the pandemic. The unemployment calls by migrants skyrocketed and increased by 120-220 percentage points more than calls by non-migrants between April and September, with no signs of abating at the end of our sample period. According to NBS, there were 291 million migrant workers in 2019, constituting 36% of the total workforce nationwide. By September 2020, the number dwindled to 179 million. The escalating number of migrants reaching out to the unemployment benefit office in our sample

[14] Note that this is not driven by females claiming unemployment benefits on behalf of their family members, which is prohibited in China.

年长的员工比年轻的员工境遇更糟:40岁及以上的员工拨打的失业电话数量比40岁以下的员工高出约20个百分点。最后，移徙者受到这一大流行病的影响最为严重。从4月到9月，移民拨打的失业电话激增，比非移民拨打的失业电话增加了120-220个百分点，在我们的样本期结束时没有减弱的迹象。根据国家统计局的数据，2019年农民工人数为2.91亿人，占全国劳动力总数的36%。到2020年9月，这一数字减少到1.79亿。在我们的样本中，向失业救济办公室求助的农民工数量不断增加，与国家统计局报告的农民工数量大幅减少相一致，这表明农民工将大规模下岗。

我们的研究结果证实了现有文献记录的最弱势的社会群体，包括移民，最容易受到不利冲击和风险的影响(Duflo和Banerjee, 2011)。最后，对女性、老年工人和移徙者造成的格外严重的影响可能反映了疫情对各行各业的不同冲击。这些群体更有可能在受到疫情严重打击的餐饮和酒店等接待行业工作，而不太可能在受影响较小的教育和高科技行业工作。

广东各城市在人口和GDP方面存在较大差异(附录表A.1)。附录图a .5的面板(a)研究了城市间异质性的影响。该数字报告了大流行治疗变量(2020年1月23日至2020年9月30日)和城市假人之间相互作用的系数，公式(2)如下。城市间的异质性相当大:虽然一些城市的失业求助增加了近70%，但少数城市未受影响，甚至出现了失业求助减少。至少有两个因素导致了城市间的差异效应。第一，城市产业构成不同。在失业电话增加最显著的七个城市中，2019年酒店和餐饮、房地产和交通运输行业的劳动力平均占比为13.9%，而受影响最小的七个城市的平均占比不到3%。为了直接说明跨行业的异质影响，我们根据方程(2)运行了一个单独的回归，其中我们将大流行治疗变量与所有13个主要行业的城市水平劳动占比相互作用。

is consistent with the massive reduction in migrant workers reported by the NBS and suggested a large-scale layoff among migrant workers.

Our results corroborate the existing literature documenting that the least advantaged social groups, including migrants, are most vulnerable to adverse shocks and risks (Duflo and Banerjee, 2011). Finally, the disproportionately more harsh impacts on females, older workers, and migrants likely reflect the pandemic's heterogeneous shocks across industries. These groups are more likely to work in hospitality industries, including restaurants and hotels, which have been hard hit by the pandemic, and less likely to work in the less affected education and high-tech industries.

There is considerable variation across cities in Guangdong in terms of population and GDP (Appendix Table A.1). Panel (a) of Appendix Figure A.5 examines the impact of heterogeneity across cities. The figure reports coefficients on the interactions of the pandemic treatment variable (which is one from January 23, 2020, to September 30, 2020) and city dummies, following equation (2). Heterogeneity across cities is sizeable: while some cities saw unemployment calls increasing by nearly 70%, a few cities were unscathed or even experienced reductions in unemployment calls. At least two factors drive the differential effects across cities. First, cities have different industry compositions. Among the seven cities that experienced the most significant increase in unemployment calls, the average share of the workforce in hotel and catering, real estate, and transportation was 13.9% in 2019, while the average share was less than 3% among the seven least affected cities. To illustrate the heterogeneous impact across industries directly, we run a separate regression following equation (2), where we interact the pandemic treatment variable with city-level labor shares by the industry for all thirteen major industries.

Panel (b) of Appendix Figure A.5 reports coefficients for all industries. The hotel and catering, real estate, and transportation sectors experienced the largest increase in unemployment calls, *ceteris paribus*. In comparison, the finance, health care, and education sectors witnessed *reductions* in unemployment calls after the lockdown in January, consistent with findings using data from other countries (Adams-Prassl et al., 2020; Alon et al., 2020). To evaluate the importance of industry compositions, we predict the number

附录图A.5的面板(b)报告了所有行业的系数。除其他因素外，酒店、餐饮、房地产和交通行业的失业电话增加最多。相比之下，金融、卫生保健和教育部门在1月封锁后的失业电话减少了，这与使用其他国家数据的发现一致(Adams-Prassl等人，2020年;阿隆等人，2020年)。为了评估行业组成的重要性，我们使用系数估计和每个城市跨行业观察到的劳动占比来预测对数(因变量)的失业呼叫数，并将预测值的范围与因变量的观察范围进行比较。城市间产业构成的变化对失业率变化的贡献率为36.3%。

of unemployment calls in logs (the dependent variable) using coefficient estimates and each city's observed labor share across industries and compare the range of predicted values with the observed range of the dependent variable. Variation in industry composition across cities contributes to 36.3% of changes in the unemployment rate.

In addition to differences in industrial composition, cities also have differential trade exposure measured by the total export relative to local GDP in 2019. For the 21 cities in Guangdong, the median export-to-GDP share in 2019 is 14.7%. Shantou city has the least exposure to international trade, whose export-to-GDP ratio is only 2.5%. At the other extreme is Dongguan, whose export-to-GDP ratio is 91.0%. As shown in panel (a) of Appendix Figure A.5, the pandemic's impact on Shantou's unemployment calls is not statistically significant from zero, while Dongguan is among the worst-hit cities, with its unemployment calls increasing 56% since the onset of the pandemic. In Appendix Table A.2, we interact the pandemic treatment variable with a city's export-to-GDP share. As expected, the interaction coefficient is statistically significant and positive. A one percentage point increase in 2019's export-to-GDP ratio is associated with a 0.19% increase in the number of unemployment calls for a given city. Similar to industry compositions, variation in the export-to-GDP ratio is also critical and explains 30% of the heterogeneity in the pandemic's unemployment impact across cities.

The sizeable estimates in Table 1 (a 27% increase in the unemployment rate) and the significant heterogeneity across cities and industries as highlighted in Appendix Figure A.5 speak to the severity of the pandemic's labor market implications, the uneven impact of the pandemic, and the importance of conducting analysis at granular levels. In addition, these results illustrate the rippling effect of the pandemic across cities within a country and countries across the world through the supply chain and the trade channels, where the industry composition of a city (or a country), the nature of the supply chain, and exposure to trade are crucial factors in determining the effect that the pandemic has on its economy (Goldberg, 2020; WTO, 2020a; WTO, 2020b; Forsythe, 2020; Von Gaudecker, 2020).

除了产业构成的差异外，以2019年出口总额相对于当地GDP衡量的城市也存在不同的贸易敞口。广东21个城市2019年出口占gdp的比例中值为14.7%。汕头市与国际贸易的接触最少，其出口占gdp的比例仅为2.5%。另一个极端是东莞，其出口占gdp的比例为91.0%。如附录图a .5 (a)组所示，疫情对汕头失业求助的影响从统计上看从零开始没有显著意义，而东莞是受疫情影响最严重的城市之一，自疫情爆发以来，其失业求助增加了56%。在附录表a .2中，我们将大流行治疗变量与城市出口占gdp的份额相互作用。如预期的那样，相互作用系数具有统计学意义且为正。2019年出口占gdp的比例每增加一个百分点，特定城市的失业人数就会增加0.19%。与产业构成类似，出口占国内生产总值比率的变化也很关键，并解释了疫情在各城市造成失业影响的30%的异质性。

表1中大量的估计数字(失业率上升了27%)以及附录图a .5所强调的不同城市和行业之间的显著差异，说明了大流行对劳动力市场影响的严重性、大流行的不均匀影响以及在细粒度层面进行分析的重要性。此外，这些结果说明了大流行通过供应链和贸易渠道在一个国家的城市和世界各国之间产生的涟漪效应，其中一个城市(或一个国家)的产业组成、供应链的性质和对贸易的接触是决定大流行对其经济影响的关键因素(Goldberg, 2020;世贸组织,2020年;世贸组织,2020 b;活力四射,2020;冯Gaudecker, 2020)。

有人可能会担心，拨打12333的人数之所以增加，仅仅是因为疫情后人们对失业救济的意识有所提高。然而，如上所示，拨打12333的人数在不同行业的变化极不均匀。城市间也存在很大的异质性，从阳江的-9%到广州的70%不等，这密切反映了城市间的行业和工人构成。这些模式不太可能纯粹是由于对疫情后失业热线的认识大幅提高所致，因为有关这些政府服务的信息主要是在国家和省一级传播的，而不是在行业(或城市)一级。

One might be concerned that increases in the number of individuals calling 12333 are merely driven by a higher awareness of unemployment benefits post the pandemic. However, as shown above, changes in the number of individuals calling 12333 are highly uneven across industries. There is also a great deal of heterogeneity across cities, and the numbers vary from -9% in Yangjiang to 70% in Guangzhou, which closely mirrors the industry and worker composition across cities. These patterns are unlikely to be purely driven by a significant increase in awareness of the unemployment hotline post the pandemic, as information on these government services is primarily disseminated at the national and provincial level, instead of at the industry (or city) level.

### *4.3 Effect on commuting patterns*

现在我们转向第二组回归分析，利用100万随机选择的用户的位置跟踪数据，利用通勤模式的变化。如果一个人在特定的时间窗口内至少访问一次他的工作地点，我们就把他视为通勤上班。为了考虑封城期间和封城后(部分)远程办公的可能性，我们使用不同的时间窗口构建了三个通勤者定义:一周、两周和一个月。例如，在以一个月为相关时间窗口的定义下，如果一个人在一个月里至少访问一个工作地点一次，他就被归类为通勤者

We now turn to the second set of the regression analysis that exploits variation in commuting patterns, using the location tracking data for one million randomly selected users. We treat an individual as commuting to work for a given time window if he visits his work location *at least once* during that time window. To accommodate the possibility of (partial) telecommuting during and after the lockdown, we have constructed three commuter definitions using different time windows: a week, two weeks, and a month. For example, under the definition that uses a month as the relevant time window, an individual is classified as a commuter for a given month if he visits a work location at least once in that month.[15]

We present three pieces of evidence that our non-commuter definition – not visiting the workplace at all during an extended period of time (such as a week, two weeks, or a month) – reflects individuals' unemployment status and that changes in non-commuters are unlikely to be primarily driven by telecommuting.

我们提出了三个证据来证明我们的不通勤定义——在很长一段时间内(如一周、两周或一个月)根本不去工作场所——反映了个人的失业状况，而不通勤的变化不太可能主要是由远程办公驱动的。

First, we examine the usage of all 51 virtual meeting apps and email apps that are available in the Apple store and Andriod gallery among commuters and non-commuters.[16] Using the two-week window to define commuters, the shares of commuters using these apps at least once in a fortnight were 30.3%, 26.1%,

[15] It is worth noting that the commuter definition throughout this section of regression analysis is different from that in Figure 2, which plots the daily number of individuals visiting their workplace during our sample period.
[16] The list of the Apps is presented in the appendix.

首先，我们研究了苹果商店和安卓库中所有51个虚拟会议应用程序和电子邮件应用程序在通勤者和非通勤者中的使用情况。16使用两周窗口来定义通勤者，在封城、第一阶段和第二阶段重新开放期间，两周至少使用一次这些应用程序的通勤者比例分别为30.3%、26.1%、24.7%。相比之下，在封城、第一阶段和第二阶段重新开放期间，两周内至少使用一次虚拟会议和电子邮件应用程序的非通勤者的比例分别为5.2%、0.6%和0.06%。当我们a)限制使用虚拟会议应用程序，b)使用一周或一个月作为定义通勤者和非通勤者的相关时间窗口时，模式非常相似。此外，通勤者和非通勤者之间使用模式的鲜明对比在2019年非常相似:通勤者每两周至少使用一次这些应用程序的比例为21.2%，而非通勤者的比例为1.0%。如果我们对2020年不通勤的衡量主要是由2020年远程办公的员工比例相对于2019年的显著增加所驱动的，那么我们预计虚拟会议应用程序的使用模式在这两年将非常不同。我们还预计，到2020年，不通勤的人使用虚拟应用的比例会更高。这两种预测都没有数据支持。

24.7% during the lockdown, Phase I, and Phase II reopening periods, respectively. In contrast, the shares of non-commuters using any virtual meetings and email apps at least once in a fortnight were 5.2%, 0.6%, and 0.06% during the lockdown, Phase I, and Phase II reopening periods, respectively. The patterns are very similar when we a) limit to virtual meeting apps, and b) use one week or one month as the relevant time window to define commuters and non-commuters. Furthermore, the sharp contrast in usage patterns between commuters and non-commuters is very similar in 2019: the share of commuters using these apps at least once in a fortnight was 21.2%, relative to 1.0% among non-commuters. If our measure of non-commuters in 2020 is primarily driven by a significant increase in the fraction of workers who telecommute in 2020 relative to 2019, we would expect the virtual meeting app usage patterns to be very different over these two years. We would also anticipate a much higher usage of virtual apps among non-commuters in 2020. Neither prediction is supported by data.

Second, our three commuters (non-commuters) measures have a remarkably high correlation (exceeding 0.92). More than 94% of individuals who are non-commuters over two weeks remain non-commuters over the entire month. These patterns hold in *both* 2019 and 2020. Suppose non-commuters in 2020 mostly consist of people who telecommute and visit offices once in a while. In that case, we should anticipate the persistence in non-commuting patterns to be lower in 2020 than that in 2019, in contrast to what we observe. These patterns provide evidence that our commuter measures accommodate flexible work modes (e.g., working at home from time to time). When individuals stop visiting their workplace over an extended period, as defined in our analysis, they are most likely not working rather than working at home.

The third piece of evidence is that the regression results and predicted unemployment are robust to the three different commuter definitions. To conserve space, we present the results that use the two-week window to define commuters in the main text. Results for the one-week and one-month commuting window and the event study that is analogous to Figure 3 but is based on commuting patterns are discussed in the appendix. We aggregate the number of commuters to the neighborhood-fortnight level, with a total of 34,965 observations.

其次，我们的三个通勤者(非通勤者)指标具有非常高的相关性(超过0.92)。超过94%的两周不上班的人在整个月都不上班。这种模式在2019年和2020年都成立。假设2020年不通勤的人大多是远程办公和偶尔去办公室的人。在这种情况下，与我们观察到的情况相反，我们应该预计2020年非通勤模式的持久性将低于2019年。这些模式提供了证据，证明我们的通勤措施适应了灵活的工作模式(例如，不时在家工作)。在我们的分析中，当一个人在很长一段时间内不去工作地点时，他们很可能不是在工作，而是在家里工作。

第三个证据是，回归结果和预测失业对三种不同的通勤定义是稳健的。为了节省篇幅，我们展示了在正文中使用两周窗口定义通勤者的结果。一周和一个月通勤窗口的结果和事件研究(类似于图3，但基于通勤模式)的结果在附录中进行了讨论。我们将通勤者的数量聚集到社区两周的水平，总共有34,965个观察结果。

表2报告了根据公式(1)对通勤者减少百分比的参数估计，再次将十天的间隔分为四个时间段。在封城期间、第一期重新开放和第二期重新开放期间，通勤人数分别减少了29%、8%和4%。在第二阶段重新开放期间减少了4个百分点，相当于每两周减少24,299名通勤者。2019年，停止通勤上班至少两周且没有新工作地点的平均人数为38729人。2020年第二阶段将减少额外的24,299名通勤者，这意味着与2019年的水平相比，停止通勤的人数将增加63%。考虑到一些企业在封城后需要一段时间才能恢复运营，我们将63%的增幅解读为疫情对广东省失业率影响的上限。

Table 2 reports parameter estimates for the percentage reduction in commuters following equation (1), again grouping the ten-day intervals into four periods. The number of commuters reduced by 29%, 8%, and 4%, during the lockdown, Phase I reopening, and Phase II reopening, respectively. The four percentage point reduction during the Phase II reopening amounts to 24,299 fewer commuters per fortnight relative to the pre-pandemic level. In 2019, the average number of individuals who stopped commuting to their work location for at least a fortnight and had no new job location was 38,729. The reduction of additional 24,299 commuters in Phase II of 2020 implies a 63% increase in the number of individuals who cease commuting relative to the 2019 level. Given that some firms take time to resume operation post the lockdown, we interpret the 63% increase as an upper bound for the pandemic's impact on unemployment in Guangdong province.

我们的结果在不同窗口长度下都非常稳定，例如一周窗口或一个月窗口(见附录):使用一周窗口时失业率增加62%，使用一个月窗口时失业率增加63%。这些模式证实了上述证据，即远程办公不太可能推动我们的研究结果，我们对通勤的衡量适应了混合工作模式。

Our results are remarkably stable across alternative window lengths, such as a one-week window or a one-month window (see appendix): the unemployment rate increased 62% using the one-week window and 63% with the one-month window. These patterns corroborate the evidence above that telecommuting is unlikely to be driving our results and that our measures of commuting accommodate hybrid work modes.

Using commuting patterns to measure unemployment has a significant advantage over measures derived from applications for unemployment benefits: it is not subject to participation bias (eligible people do not participate). Due to the inertia, lack of information, stigma, time, and "hassle" associated with applications, it is estimated that 66% of eligible households do not participate in major social programs in the U.S. (Ribar, 2020). Non-participation is much more severe in developing countries due to the limited program benefits. As a result, commuting patterns observed over an extended period of time provide a real-time and likely more accurate indicator of the underlying labor market dynamics.

Column 2 of Table 2 examines changes in work duration among individuals working on-site. Hours on-site dropped by 21% during the lockdown period but returned to the 2019 level in phase II reopening. The pandemic does not seem to have brought about dramatic changes in the nature of working on-site, lending further support to our strategy of measuring unemployment based on changes in individuals

使用通勤模式来衡量失业率，比从失业救济申请中得出的方法有显著的优势:它不受参与偏差的影响(符合条件的人不参与)。由于与申请相关的惯性、信息缺乏、耻辱、时间和"麻烦"，据估计，66%的符合条件的家庭没有参加美国的主要社会项目(Ribar, 2020)。由于项目收益有限，发展中国家的不参与情况要严重得多。因此，在一段较长时间内观察到的通勤模式提供了一个实时的、可能更准确的潜在劳动力市场动态指标。

表2第2列审查了在现场工作的个人工作时间的变化。在封锁期间，现场工作时间减少了21%，但在第二阶段重新开放时恢复到了2019年的水平。大流行似乎没有给现场工作的性质带来巨大变化，这进一步支持了我们根据通勤上下班人员的变化来衡量失业的战略。附录数据要求寄出。8和附录A.3展示了跨人口统计群体、城市和行业的异质影响。调查结果与基于失业电话的调查结果在性质上类似:自大流行以来，女性、40岁以上的工人和移徙工人通勤和就业的减少更为明显。城市和行业之间也存在大量的异质性，与图a .5中报告的模式密切相关。最后，在大流行期间，出口占gdp比例较大的城市通勤人数减少幅度较大。

commuting to work. Appendix Figures A.6-A.8 and Appendix Table A.3 present heterogeneous impacts across demographic groups, cities, and industries. The results are qualitatively similar to those that are based on unemployment calls: females, workers over 40, and migrant workers experienced a more pronounced reduction in commuting, and hence employment, since the pandemic. There is also a significant amount of heterogeneity across cities and industries, closely mirroring patterns reported in Figure A.5. Finally, cities with a larger export-to-GDP ratio experience a larger reduction in the number of commuters during the pandemic.

## 5. Conclusion

Our analysis contributes to the recent studies on the broad labor market impacts of the pandemic (Bick and Blandin, 2020; Coibion, 2020; Ahn and Hamilton, 2020; Cajner et al., 2020). Compared to these studies, our mobile phone data benefit from large sample size and fine resolutions in both temporal and spatial dimensions. The pandemic's estimated impact on unemployment (27-63% increase) in China is much smaller than that in the United States, partly due to the differences in the composition of the economy between these two countries. The service sector, which has been hard hit by the pandemic, employs 79% of the workforce and produces 68% of the GDP in the United States, compared to 47% and 50% in China in 2018. In addition, the draconian measures adopted in China to control the pandemic have reduced the spread of the virus more effectively (Hsiang et al., 2020; Kraemer et al., 2020; Li et al., 2020; Zhang et al., 2020) and likely mitigated the impact on the economy as a result.

Our findings on the uneven labor market impacts across demographic groups and industries are consistent with recent studies in other countries (Adams-Prassl et al., 2020; Alon et al., 2020*)*. However, we conduct the first quantitative analysis on China using large-scale datasets. Our research adds to the literature by showing that the pandemic's adverse impact on the labor market is more severe in areas that rely more heavily on export and hence more exposed to external shocks through global trade channels.

我们的分析有助于最近关于大流行对劳动力市场的广泛影响的研究(比克和布兰德，2020年;Coibion, 2020;安和汉密尔顿，2020年;Cajner等人，2020年)。与这些研究相比，我们的手机数据得益于大样本量和高分辨率的时间和空间维度。据估计，疫情对中国失业率(增长27%至63%)的影响远小于美国，部分原因是两国经济构成的差异。2018年，美国受疫情影响严重的服务业雇佣了79%的劳动力，生产了68%的GDP，而中国的这一比例分别为47%和50%。此外，中国为控制大流行而采取的严厉措施更有效地减少了病毒的传播(Hsiang等人，2020;Kraemer等人，2020年;李等人，2020;Zhang等人，2020年)，结果可能减轻了对经济的影响。

我们关于跨人口统计群体和行业的不均衡劳动力市场影响的发现与其他国家最近的研究一致(Adams-Prassl等人，2020年;阿隆等人，2020年)。然而，我们首次使用大规模数据集对中国进行定量分析。我们的研究补充了文献，表明大流行对劳动力市场的不利影响在严重依赖出口、因此更容易通过全球贸易渠道受到外部冲击的地区更为严重。

**References**


Adams-Prassl, A., Boneva, T., Golin, M., & Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics*, 189, 104245.

Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2010). Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data. *Transportation Research Part C: Emerging Technologies*, 18(1), 45-54.

Ahn, H. J., & Hamilton, J. D. (2021). Measuring labor-force participation and the incidence and duration of unemployment. *Review of Economic Dynamics*.

Alon, T. M., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020). The impact of COVID-19 on gender equality (No. w26947). National Bureau of Economic Research.

Atkin, D., Chen, K., & Popov, A. (2020). The returns to serendipity: Knowledge spillovers in Silicon Valley. Unpublished working paper.

Barwick, P. J., Liu, Y., Patacchini, E., & Wu, Q. (2019). Information, mobile communication, and referral effects (No. w25873). National Bureau of Economic Research.

Bick, A., & Blandin, A. (2020). Real-time labor market estimates during the 2020 coronavirus outbreak. Available at SSRN 3692425.

Cai, F., Du, Y., & Wang, M. (2013). Demystify the labor statistics in China. *China Economic Journal*, 6(2-3), 123-133.

Cajner, T., Crane, L. D., Decker, R. A., Grigsby, J., Hamins-Puertolas, A., Hurst, E., ... & Yildirmaz, A. (2020). The US labor market during the beginning of the pandemic recession (No. w27159). National Bureau of Economic Research.

Card, D. (2011). Origins of the unemployment rate: the lasting legacy of measurement without theory. *American Economic Review*, 101(3), 552-57.

Chen, B. Y., Wang, Y., Wang, D., Li, Q., Lam, W. H., & Shaw, S. L. (2018). Understanding the impacts of human mobility on accessibility using massive mobile phone tracking data. *Annals of the American Association of Geographers*, 108(4), 1115-1133.

Chen, M. K., Haggag, K., Pope, D. G., & Rohla, R. (2019). Racial disparities in voting wait times: evidence from smartphone data. *Review of Economics and Statistics*, 1-27.

Chen, M. K., & Pope, D. G. (2020). Geographic mobility in America: Evidence from cell phone data (No. w27072). National Bureau of Economic Research.

Chevalier, J. A., Schwartz, J. L., Su, Y., & Williams, K. R. (2021). Measuring Movement and Social Contact with Smartphone Data: A Real-time Application to COVID-19 (No. 2280). Cowles Foundation for Research in Economics, Yale University.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). Labor markets during the COVID-19 crisis: A preliminary view (No. w27017). National Bureau of Economic Research.

Couronné, T., Smoreda, Z., & Olteanu, A. M. (2013). Chatty mobiles: Individual mobility and communication patterns. arXiv preprint arXiv:1301.6553.

Duflo, E., & Banerjee, A. (2011). Poor economics. PublicAffairs.

Feng, S., & Hu, Y. (2013). Misclassification errors and the underestimation of the US unemployment rate. *American Economic Review*, 103(2), 1054-70.

Feng, S., Hu, Y., & Moffitt, R. (2017). Long run trends in unemployment and labor force participation in urban China. *Journal of Comparative Economics*, 45(2), 304-324.

Forsythe, E. (2020). Labor market flows and the Covid-19 economy. Available at SSRN 3586074.

Giles, J., Albert, P., & Zhang, J. (2005). What is China's true unemployment rate?. *China Economic Review*, 16(2), 149-170.

Goldberg, P. (2020). Global value chains, COVID-19, and the future of trade. Paper presented at Princeton University.

Gong, H., Hassink, R., Tan, J., & Huang, D. (2020). Regional Resilience in Times of a Pandemic Crisis: The Case of COVID-19 in China. *Tijdschrift voor economische en sociale geografie*, 111(3), 497-512.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779-782.

Gupta, S., Nguyen, T. D., Rojas, F. L., Raman, S., Lee, B., Bento, A., ... & Wing, C. (2020). Tracking public and private responses to the COVID-19 epidemic: evidence from state and local government actions (No. w27027). National Bureau of Economic Research.

Heffetz, O., & Reeves, D. B. (2019). Difficulty of reaching respondents and nonresponse Bias: Evidence from large government surveys. *Review of Economics and Statistics*, 101(1), 176-191.

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., ... & Wu, T. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584(7820), 262-267.

Jones, S. R., & Riddell, W. C. (1999). The measurement of unemployment: An empirical approach. *Econometrica*, 147-161.

Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., ... & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493-497.

Kreindler, G. E., & Miyauchi, Y. (2021). Measuring commuting and economic activity inside cities with cell phone records (No. w28516). National Bureau of Economic Research.

Krueger, A. B., Mas, A., & Niu, X. (2017). The evolution of rotation group bias: Will the real unemployment rate please stand up?. *Review of Economics and Statistics*, 99(2), 258-264.

Liu, Q. (2012). Unemployment and labor force participation in urban China. *China Economic Review*, 23(1), 18-33.

Meyer, B. D., Mok, W. K., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4), 199-226.

Poterba, J. M., & Summers, L. H. (1986). Reporting errors and labor market dynamics. *Econometrica*, 1319-1338.

Ribar, D. C. (2014). How to improve participation in social assistance programs. *IZA World of Labor*.

Tao, Y., & Yuan, Y. (Eds.). (2016). Annual Report on the Development of China's Special Economic Zones (2016): Blue Book of China's Special Economic Zones. Springer.

Toole, J. L., Lin, Y. R., Muehlegger, E., Shoag, D., González, M. C., & Lazer, D. (2015). Tracking employment shocks using mobile phone data. *Journal of The Royal Society Interface*, 12(107), 20150185.

von Gaudecker, H. M., Holler, R., Janys, L., Siflinger, B., & Zimpelmann, C. (2020). Labour Supply in the Early Stages of the COVID-19 Pandemic: Empirical Evidence on Hours, Home Office, and Expectations. IZA Working Paper.

World Trade Organization. (2020a). Trade shows signs of rebound from COVID-19, recovery still uncertain. (WTO Technical Report Series No. 862).

World Trade Organization. (2020b). The economic impact of COVID-19 on women in vulnerable sectors and economies. (WTO Technical Report).

Yu X., & Zhang C. (2010). Reducing inequality for shared growth in China: strategy and policy options for Guangdong province. The World Bank report.

Zhang, R., Li, Y., Zhang, A. L., Wang, Y., & Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences*, 117(26), 14857-14863.

**Figure 1. Differences in the number of individuals calling 12333 between 2019 and 2020**

**Notes:** This graph depicts the daily number of individuals calling the unemployment-benefit hotline 12333 (in thousands) in 2020 minus that in 2019. The red and blue bars (left y-axis) denote daily differences in the number of individuals calling 12333 between 2020 and 2019. The dark dash curve (the right y-axis) illustrates the cumulative difference in the number of individuals calling 12333 since 60 days before the event date. The event days are based on the lunar calendar. For 2020, the event date (or day zero) is the lockdown, January 23, 2020, two days before the Chinese New Year. Correspondingly, the event date in 2019 is February 3, 2019, also two days before the 2019 Chinese New Year. Phase I reopening started on February 24, 2020, 32 days after the lockdown, when people were allowed to go back to work and visit outdoor public places. Phase II reopening, or full reopening, started on May 9, 2020, 108 days after the lockdown. Shopping malls, supermarkets, restaurants were allowed to fully reopen.

**Figure 2. Differences in daily workers on-site and hours on-site between 2019 and 2020**

(a) Differences in daily workers on-site between 2020 and 2019

(b) Differences in daily work hours among individuals working on-site between 2020 and 2019

**Notes**: Panel (a) depicts the daily number of individuals working on-site in 2020 minus that in 2019, and panel (b) depicts the difference in daily hours working on-site, conditioning on being on-site. The event days are based on the lunar calendar. For 2020, the event date (or day zero) is the lockdown, January 23, 2020, two days before the Chinese New Year. Correspondingly, the event date in 2019 is February 3, 2019, also two days before the 2019 Chinese New Year. Phase I reopening started on February 24, 2020, 32 days after the lockdown, when people were allowed to go back to work and visit outdoor public places. Phase II reopening, or full reopening, started on May 9, 2020, 108 days after the lockdown. Shopping malls, supermarkets, restaurants were allowed to fully reopen.

**Figure 3. Event study using unemployment calls**

(a) All

(b) Female v.s. male

(c) 40 and above v.s. under 40 years old

(d) Migrants v.s. non-migrants

**Notes:** All event-study graphs plot coefficients $\beta_q$ from equation (1), which are the percentage changes in individuals making unemployment calls in 2020 relative to 2019. The dependent variable is the log number of individuals calling the unemployment-benefit hotline 12333 in panel (a), the difference between female and male calling the hotline (in logarithm) in panel (b), the difference between individuals age 40 and above and those below 40-year-old making unemployment calls (in logarithm) in panel (c), and the difference between non-migrants and migrants making unemployment calls (in logarithm) in panel (d). All regressions include neighborhood, day-of-week, event-day, holiday, and the treatment group fixed effects. The standard errors are clustered at the event-day level.

The event days are based on the lunar calendar. For 2020, the event date (or day zero) is the lockdown, January 23, 2020, two days before the Chinese New Year. Correspondingly, the event date in 2019 is February 3, 2019, also two days before the 2019 Chinese New Year. Phase I reopening started on February 24, 2020, 32 days after the lockdown, when people were allowed to go back to work and visit outdoor public places. Phase II reopening, or full reopening, started on May 9, 2020, 108 days after the lockdown. Shopping malls, supermarkets, restaurants were allowed to fully reopen.

**Table 1. Effects on the number of individuals making unemployment calls and call duration**

| VARIABLES | (1)<br>No. of individuals making calls (in log) | (2)<br>Duration per call (in log) |
|---|---|---|
| **Panel A:** | | |
| 1-30 days before lockdown | 0.03 | 0.03 |
| | (0.03) | (0.04) |
| Lockdown period | -0.37*** | -0.36*** |
| | (0.06) | (0.08) |
| Phase I re-opening | 0.25*** | 0.24*** |
| | (0.03) | (0.05) |
| Phase II re-opening | 0.45*** | 0.56*** |
| | (0.02) | (0.04) |
| **Panel B:** | | |
| Pandemic period | 0.27*** | 0.33*** |
| (Lockdown + Phase I + Phase II) | (0.03) | (0.03) |
| Observations | 489,514 | 489,514 |
| R-squared | 0.81 | 0.57 |
| Neighborhood FE | Yes | Yes |
| Event-day FE | Yes | Yes |
| Day-of-week FE | Yes | Yes |
| Holidays FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** Panel A in this table examines the percentage change in the number of individuals making unemployment calls and call duration as a result of the pandemic following equation (1), except that the ten-day intervals are grouped into four periods: before lockdown, during the lockdown, Phase I reopening, and Phase II full reopening. In panel B, we combine the three periods during the lockdown, Phase I reopening, and Phase II full reopening into one dummy variable, named the 'Pandemic period,' to examine the average impact of the pandemic. The observations are at the neighborhood by day level. The dependent variables in columns (1)-(2) are the number of individuals making unemployment calls (in logarithm) and the average duration of unemployment calls (in seconds, in logarithm), respectively. Both columns include neighborhood, day-of-week, event-day, holiday, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at the event day. $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

**Table 2. Effects on commuters and working hours**

| VARIABLES | (1)<br>No. of commuters (in log) | (2)<br>Working hours (in log) |
|---|---|---|
| 1-30 days before lockdown (=1) | 0.03 | 0.01 |
| | (0.02) | (0.01) |
| Lockdown period (=1) | -0.29*** | -0.21*** |
| | (0.03) | (0.02) |
| Phase I re-opening (=1) | -0.08*** | -0.08*** |
| | (0.03) | (0.01) |
| Phase II re-opening (=1) | -0.04** | -0.02 |
| | (0.02) | (0.02) |
| Observations | 34,965 | 34,965 |
| R-squared | 0.98 | 0.95 |
| Neighborhood FE | Yes | Yes |
| Event-fortnight FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table examines the percentage change in the number of commuters and duration of on-site working hours as a result of the pandemic, following equation (1), except that the ten-day intervals are grouped into four periods: before lockdown, during the lockdown, Phase I reopening, and Phase II full reopening. The observations are at the neighborhood by fortnight level. The dependent variables in columns (1)-(2) are the number of commuters (in logarithm) and average working hours for commuters (in logarithm), respectively. A commuter is someone who visits his work location at least once in a given two-week window. Those who do not visit their work location at all during that two-week window are coded as non-commuters. Both columns include neighborhood, event-fortnight, holiday, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at the event-fortnight level. $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

**Online Appendix for**

**"COVID-19 Pandemic and Unemployment: Evidence from Mobile Phone Data in China"**

Teng Li, Panle Jia Barwick, Yongheng Deng, Xinfei Huang, Shanjun Li

## Data description

China's mobile phone penetration is high among developing countries. Figure A.1 shows a strong correlation between the number of China Mobile users and the number of residents by city. Cities with a higher GDP per capita (represented by the size of the circles in Figure A.1) tend to have higher mobile phone ownership.

Our data come from China Mobile, the dominant cellular service provider in China. We use the number of individuals calling 12333 instead of the number of calls to 12333 to measure the number of individuals claiming unemployment benefits, as people might make multiple calls to claim their benefits. Calls that failed to go through to the receiving party and calls shorter than 30 seconds are excluded from the analysis.

An important feature of our mobile phone data is the location tracking information (in longitude and latitude) that is collected by one million randomly selected mobile devices during 5-minute intervals from January 2018 to September 2020. Recent developments and the widespread diffusion of geospatial data acquisition technologies have enabled the creation of highly accurate spatial and temporal data (Gonzalez et al., 2008). Note that passive collection of geolocation information – which underlies our data collection procedure -- works on all traditional mobile networks (2G, 3G, or 4G).

To construct users' work locations, we first identify the location where the user spent the most time between 9 am and 6 pm during weekdays for each month. If an individual spends at least 5 hours a day between 9 am and 6 pm for at least fifteen workdays at his primary location for a given month, we code this location as this individual's work location for that month. For each individual and a given month, we define his home location as the location that has the most frequent usage between 10 pm and 7 am. Phone usage during 7 am-9 am and 6 pm-10 pm is excluded because people are likely on the move during these time intervals. All individuals have a home location.

We provide two pieces of evidence that our assignment of home and work locations captures an intuitive spatial distribution of users in our sample. First, we use the coordinates of work and residential locations to compute the commuting distance for users with valid job location information. The distribution of commuting distance decays exponentially (Figure A.9), which is intuitive. Additionally, the average commuting distance in our sample period is around 6.6km, close to the average commuting distance of 6.9km reported in the 2017 travel survey by the Guangzhou Municipal Transportation Bureau (GMTB).[17] Second, for the city of Guangzhou (the provincial capital city), Figure A.10 plots the log difference between the number of users at 11 am and the number of users at 11 pm, averaged separately for weekdays and weekends in 2019. The figure includes all geographic locations recorded in the data. On both weekdays and weekends, the center of the city gains population, and the suburbs lose population during the daytime relative to the nighttime. But these differences are much more pronounced on weekdays than weekends, especially in the center of the city. For instance, the enlarged area is a famous industrial park in Guangzhou, which clearly shows that the daytime population is much larger than the nighttime population. The differences are amplified on weekdays. The spatial pattern of population density is consistent with those from the report by GMTB.

**Guangdong's lockdown**

Guangdong's provincial government acted swiftly and adopted vigilant procedures since the onset of the pandemic. Guangdong was one of the first provinces to release detailed information (frequency, location, gender, etc.) on the newly confirmed cases, starting from as early as February 3, 2020. These procedures proved successful and have kept the number of daily confirmed cases under a few handfuls since the full reopening (Figure A.2). As shown in Figure A.2, the daily confirmed new cases reached a peak of 254 on January 31 and quickly reduced to under 50 three weeks into the lockdown period. The number of cases has been modest since then and varies between 0 and 34 throughout the Phase I and Phase II reopening.

[17] Refer to the report "Comprehensive Report of New Round Traffic Survey of Guangzhou" by Guangzhou Transportation Planning Research Institute in 2017.

The lockdown in Guangdong lasted 32 days from January 23 to February 24, 2020. The provincial government issued an order on February 6, 2020, and encouraged workers in some industries to get back to work after February 24. It is worth noting that the lockdown procedures in Guangdong are not as strict as the lockdown procedures implemented in the epicenter Wuhan. On February 24, 2020, Guangdong province entered Phase I reopening, which lasted 76 days. During Phase I reopening, people were allowed (and encouraged in certain industries) to go back to work and visit outdoor public places. Phase II reopening, or full reopening, officially started on May 9, 2020, when all businesses, including shopping malls, supermarkets, and restaurants, were allowed to open fully. The only exception was movie theaters that remained closed till mid-July.

While Guangdong's number of COVID cases is low, it does not imply that the pandemic has a modest or no effect on the local economy: the measures implemented to reduce the health impact of the pandemic could significantly affect the economy. As shown in our analysis in the main text, the pandemic has inflicted sizeable damage to Guangdong's labor market, leading to a 27%-63% increase in the unemployment rate. As Guangdong's economy is among the most vigilant across provinces in China, the aggregate labor market implications could be much more severe than those suggested by the national statistics.

**Virtual Apps and Email Apps Usage**

To gauge the prevalence of telecommuting, we obtain an exhaustive list of all 21 virtual meeting apps and 30 email apps available in the Apple store and Andriod gallery. The virtual meeting apps include: Tencent meeting, DingDing, Zoom, SKYPE, uu online, Alitong online, Ailiao, Chubao, Feige, Feiyin, Laidian, Shangqitong, Shuoba, Tongtong, Weihui, Weiwei, Yiliao, Youhuatong, Youliao, Youxin, Zhangshangbao.

The email apps include: 139 mail, 139 light mail, 189 mail, 21CN mail, 21CN light mail, 263 mail, Gmail, Live mail, Microsoft Outlook, QQ mail, TOM, Ali mail, Ke space, Sohu mail, Qixinbao, Tencent mail, Wangyi mail, Woo mail, Mi mail, Sina mail, Yahoo mail, Youqia, Mail master, Yun home, China Mobile mail, Baidu mail, Hotmail, Foxmail, Coremail.

We have constructed three measures of commuters using different window lengths: one week, two weeks, and one month. For example, using the one-month window, an individual who visits his work location at least once in a given month is coded as a commuter that month, and those who do not visit their work location at all that month are coded as non-commuters. We have examined app usage among commuters and non-commuters and find the patterns to be very similar across all three definitions of commuters. Using the one-week window, the fraction of non-commuters who use these apps in a given week is 6.5%, 1.7%, and 0.07% during the lockdown, Phase I, and Phase II reopening, respectively. In contrast, the fraction of commuters who use these apps in a given week were 25.7%, 22.3%, 23.3% during the lockdown, Phase I, and Phase II reopening periods, respectively. Using the one-month window, the fraction of non-commuters who use these apps in a given month is 5.0%, 0.5%, and 0.06% during the lockdown, Phase I, and Phase II reopening, while the fraction on commuters who use these apps in a given month is 32.6%, 26.9%, 25.1%, respectively. The patterns are very similar if we limit to virtual meeting apps only.

Importantly, the sharp contrast in app usage between commuters and non-commuters is very similar in 2019 as that in 2020. Specifically, in 2019 and using the one-week window, the share of non-commuters using any of the virtual meetings and email apps at least once in a week was 1.1%. In contrast, the share of commuters using these apps at least once was 20.5%. The fraction was 1% and 21.6% for non-commuters and commuters using the one-month window, respectively.

If our measures of non-commuters in 2020 are primarily driven by a significant increase in the fraction of workers who telecommute in 2020 relative to 2019, we would expect the virtual meeting app usage patterns to be very different over these two years. We would also anticipate a much higher usage of virtual apps among non-commuters in 2020. Neither prediction is supported by data.

**Regression Analysis**

China Mobile delineates Guangdong province into 787 mobile phone cell tower areas (similar to zip codes in the U.S.) for billing purposes. All regressions control for cell-tower-area (or neighborhood) fixed effects. Regressions for unemployment calls aggregate the sample to the neighborhood and day level, with a total

of 489,514 observations. Regressions for commuting patterns further aggregate to the neighborhood-week, neighborhood-fortnight, and neighborhood-month level when appropriate.

As Table A.1 illustrates, cities in Guangdong differ substantially in terms of both population and GDP in 2019. Among the 21 cities, Guangzhou has the largest population (15.31 million), while Yunfu only has a 2.55 million population. The economic scale of the largest city, Shenzhen, at $390 billion in 2019, is almost 30 times as large as that of the least city, Yunfu.

Table A.2 uses equation (2) to explore the heterogeneous effects of trade exposure, which is measured by the ratio of a city's total export to its local GDP in 2019. The interactions between pandemic period dummy and export-to-GDP share are statistically significant and positive, indicating that cities with larger export-to-GDP shares experienced larger increases in unemployment calls.

We have repeated the event study analysis using commuting patterns and plot the event coefficients in Figure A.6. The dependent variable is the log number of workers who commute to job locations at least once during a week (panel (a) of Figure A.6) and the log number of hours worked on-site (panel (b) of Figure A.6), conditioning on commuting. The number of people who work on-site dropped as much as 40% during the lockdown, as businesses were temporarily shut down and non-essential workers were prevented from commuting to work. During Phase I reopening, the number of people working on-site was 5-25% lower than the 2019 level. When all businesses were allowed to open during Phase II reopening, the number of people working on-site is still about 4-10% lower than the 2019 level.

Figures A.7-A.8 and Table A.3 present heterogeneous impacts across demographic groups, cities, and industries. The results are qualitatively similar to those that are based on unemployment calls: females, workers over 40, and migrant workers experienced a more pronounced reduction in commuting, and hence employment, since the pandemic. There is also a significant amount of heterogeneity across cities and industries, closely mirroring patterns reported in Figure A.5. Finally, cities with a larger export-to-GDP ratio experience a larger reduction in the number of commuters during the pandemic.

The analysis of commuting patterns in the main text uses the two-week window to define a commuter. We have constructed two alternative measures of commuters using the one-week window and one-month window. It is worth noting that these three commuter measures are very highly correlated: the correlation between the one-week and two-week measure is 0.95, the correlation between the one-week and one-month measure is 0.92, and the correlation between the two-week and one-month measure is 0.97. In addition, more than 94% of individuals who are non-commuters over two weeks remain non-commuters over the entire month. This is evidence that our commuter measures accommodate flexible work modes (e.g., working at home from time to time). When individuals stop visiting their workplace over an extended period as defined in our analysis, they are essentially not working rather than working at home.

In Tables A.4 and A.5 we repeat the entire analysis using a one-week window and one-month window as alternative cutoffs. The estimates are remarkably close to those in Table 2 in the main text.

Our main analysis excludes all users under the age of 18. As some users between the age of 18 and 25 might still be in schools, we exclude users under the age of 25 as a robustness analysis. Results on the percentage changes in calls (Table A.6), and commuters (Table A.7), bearly change when we limit to users 25 and above.

**Figure A.1. Mobile users vs. residents in 2019**

**Notes:** This graph presents the relationship between average daily mobile users and residents for cities in Guangdong in 2019. The solid line is a 45-degree line. The size of each marker denotes GDP per capita in 2019. We label the city names for those whose GDP per capita in 2019 is above 60,000 RMB (around 8,500 USD). The abbreviated city names are listed in Table A.1.

**Figure A.2. Newly confirmed COVID-19 cases in Guangdong province**

**Notes:** This graph shows the number of daily newly confirmed COVID-19 cases in Guangdong province. The provincial government started to report the number of COVID-19 cases on January 10, 2020, 13 days before the lockdown in Guangdong.

**Figure A.3. Correlation of Baidu index between keywords of "12333" and "unemployment insurance" in Guangdong province**

**Notes:** This graph shows the weekly Baidu Search Index for the keywords – "12333" and "unemployment insurance" -- in Guangdong province from 2019 to 2020. Baidu Index, which is similar to Google Trends, is a keyword-analysis tool launched by Baidu, the largest search engine company in China. It reflects the search frequency of certain keywords on the Baidu website. The correlation between the two keywords during the sample period is 0.83.

**Figure A.4. Unemployment rate by city in 2019 based on the number of individuals making unemployment calls**

**Notes:** This graph shows our estimated unemployment rate, which is the ratio of individuals making unemployment calls over the size of the labor force by city in 2019. The correlation between our unemployment measure and the official unemployment measure at the city level is 0.7 in 2019.

**Figure A.5. Heterogeneity across industries and cities using unemployment calls**

(a) By city

(b) By industry

**Notes:** This figure illustrates heterogeneity across cities (panel (a)) and industries (Panel (b)) following equation (2). In panel (a), we add interactions between the after-lockdown dummy and city fixed effects. In panel (b), we add interactions between the after-lockdown dummy and a city's share of employment in each of the 13 industries. A positive change indicates an increase in unemployment relative to 2019. Both regressions include neighborhood, day-of-week, event-day, holiday, and the treatment group fixed effects. The standard errors are clustered at the event-day level.

**Figure A.6. Event study on the number of individuals commuting to work and hours on-site**

(a) No. of people commuting to work

(b) Hours working on-site

**Notes:** The event-study graphs plot coefficients $\beta_q$ from equation (1), which are the percentage changes in the number of individuals commuting to work (panel (a)), and hours of on-site conditioning on commuting (panel (b)). This analysis uses one week to define commuters. Both regressions include neighborhood, event-week, and the treatment group fixed effects. The standard errors are clustered at the event-week level. Results are qualitatively similar, using two weeks or one month to define commuters.

**Figure A.7. Heterogeneity in commuting patterns across demographic groups**

(a) Female v.s. male

(b) 40 and above v.s. under 40 years old

(c) Migrants v.s. non-migrants

**Notes:** All event-study graphs plot coefficients $\beta_q$ from equation (1), which are the percentage changes in individuals commuting to work in 2020 relative to 2019. The dependent variable is the difference between female and male commuting to work (in logarithm) in panel (a), the difference between individuals age 40 and above and those below 40-year-old commuting to work (in logarithm) in panel (b), and the difference between migrants and non-migrants commuting to work (in logarithm) in panel (c). This analysis uses one week to define commuters. All regressions include neighborhood, event-week, and the treatment group fixed effects. The standard errors are clustered at the event-week level. Results are qualitatively similar, using two weeks or one month to define commuters.

**Figure A.8. Heterogeneity across industries and cities using commuting sample**

(a) By city

(b) By industry

**Notes:** This figure illustrates heterogeneity across cities (panel (a)) and industries (Panel (b)) following equation (2). In panel (a), we add interactions between the after-lockdown dummy and city fixed effects. In panel (b), we add interactions between the after-lockdown dummy and a city's share of employment in each of the 13 industries. A negative change indicates a decrease in commuters relative to 2019. This analysis uses one week to define commuters. Both regressions include neighborhood, event-week, and the treatment group fixed effects. The standard errors are clustered at the event-week level. Results are similar using two weeks or one month to define commuters.

**Figure A.9. The number of people commuting by commuting distance**

**Notes:** This graph shows the number of people commuting at each distance. We use the coordinates of job and home locations to compute the spherical commuting distance.

**Figure A.10. Differences in (log) population during daytime and nighttime in Guangzhou**

(a) Weekdays

(b) Weekends

**Notes:** Panel (a) displays the average daily log difference of users during daytime (11 am) and nighttime (11 pm) at the centroid of every phone tower on the weekdays of 2019. In contrast, panel (b) shows the same measure on the weekends of 2019.

**Table A.1. Cities in Guangdong province**

| Cities | Population in 2019 (million) | GDP in 2019 (billion USD) |
|---|---|---|
| Shenzhen (SZ) | 13.44 | 390.25 |
| Guangzhou (GZ) | 15.31 | 342.44 |
| Foshan (FS) | 8.16 | 155.81 |
| Dongguan (DG) | 8.46 | 137.43 |
| Huizhou (HZ) | 4.88 | 60.54 |
| Zhuhai (ZH) | 2.02 | 49.8 |
| Maoming (MM) | 6.41 | 47.13 |
| Jiangmen (JM) | 4.63 | 45.6 |
| Zhongshan (ZS) | 3.38 | 44.94 |
| Zhanjiang (ZJ) | 7.36 | 44.42 |
| Shantou (ST) | 5.66 | 39.04 |
| Zhaoqing (ZQ) | 4.19 | 32.59 |
| Jieyang (JY) | 6.11 | 30.46 |
| Qingyuan (QY) | 3.89 | 24.61 |
| Shaoguan (SG) | 3.03 | 19.11 |
| Yangjiang (YJ) | 2.57 | 18.73 |
| Meizhou (MZ) | 4.38 | 17.2 |
| Chaozhou (CZ) | 2.66 | 15.67 |
| Shanwei (SW) | 3.02 | 15.66 |
| Heyuan (HY) | 3.11 | 15.65 |
| Yunfu (YF) | 2.55 | 13.36 |

**Notes:** This table shows the 2019 population and GDP for each city in Guangdong province. Data source: The Bureau of Statistics in Guangdong.

**Table A.2. Heterogeneity by export-to-GDP ratio**

| VARIABLES | (1)<br>No. of individuals making calls (in log) | (2)<br>Duration per call (in log) |
|---|---|---|
| Pandemic period * Export/GDP in 2019 (%) | 0.0019*** | 0.0019*** |
| | (0.0003) | (0.0004) |
| Pandemic period (=1) | 0.2086*** | 0.2688*** |
| | (0.0232) | (0.0381) |
| Observations | 489,514 | 489,514 |
| R-squared | 0.80 | 0.57 |
| Neighborhood FE | Yes | Yes |
| Event-day FE | Yes | Yes |
| Day-of-week FE | Yes | Yes |
| Holidays FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table examines whether the pandemic's effect differs across cities with varying exposure to international trade following equation (2), where we interact the pandemic treatment with a city's 2019 export-to-GDP ratio (%). The dependent variables in columns (1)-(2) are the number of unemployment calls (in logarithm) and average duration of unemployment calls in seconds (in logarithm), respectively. The observations are at the neighborhood and day level. Both columns include neighborhood, day-of-week, event-day, holiday, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at the event-day. $*$ p <0.1, $**$ p <0.05, $***$ p < 0.01.

**Table A.3. Heterogeneity by export-to-GDP ratio using commuting patterns**

| VARIABLES | (1)<br>No. of commuters (in log) | (2)<br>Working hours (in log) |
|---|---|---|
| Pandemic period * Export/GDP in 2019 (%) | -0.0026*** | -0.0002* |
| | (0.0002) | (0.0001) |
| Pandemic period (=1) | -0.0635*** | -0.0623*** |
| | (0.0134) | (0.0065) |
| Observations | 34,965 | 34,965 |
| R-squared | 0.98 | 0.81 |
| Neighborhood FE | Yes | Yes |
| Event-fortnight FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table examines whether the pandemic's effect differs across cities with varying exposure to international trade following equation (2), where we interact the pandemic treatment with a city's 2019 export-to-GDP ratio. The dependent variables in columns (1)-(2) are the number of commuters (in logarithm) and average working hours for commuters (in logarithm), respectively. An individual who visits his work location at least once in a given fortnight is coded as a commuter, and an individual who does not visit his work location at all during that fortnight is coded as a non-commuter. Both columns include neighborhood, event-fortnight, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at event-fortnight. $*$ p <0.1, $**$ p < 0.05, $***$ p < 0.01. Results are similar using one week or one month to define commuters.

**Table A.4. Effect on commuters and working hours (one-week window)**

| VARIABLES | (1)<br>No. of commuters (in log) | (2)<br>Working hours (in log) |
|---|---|---|
| 1-30 days before lockdown (=1) | -0.01 | -0.00 |
| | (0.02) | (0.01) |
| Lockdown period (=1) | -0.27*** | -0.20*** |
| | (0.05) | (0.02) |
| Phase I re-opening (=1) | -0.10*** | -0.08*** |
| | (0.03) | (0.02) |
| Phase II re-opening (=1) | -0.05*** | -0.02 |
| | (0.01) | (0.02) |
| Observations | 69,930 | 69,930 |
| R-squared | 0.97 | 0.94 |
| Neighborhood FE | Yes | Yes |
| Event-week FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table replicates the analysis in Table 2, except that the data are aggregated at the neighborhood-week level. The dependent variables in columns (1)-(2) are the number of commuters (in logarithm) and average working hours for commuters (in logarithm), respectively. An individual who visits his work location at least once in a given week is coded as a commuter, and an individual who does not visit his work location at all during that week is coded as a non-commuter. Both columns include neighborhood, event-week, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at event-week. $*$ p <0.1, $**$ p < 0.05, $***$ p < 0.01.

**Table A.5. Effect on commuters and working hours (one-month window)**

| VARIABLES | (1)<br>No. of commuters (in log) | (2)<br>Working hours (in log) |
|---|---|---|
| 1-30 days before lockdown (=1) | 0.03 | 0.01 |
| | (0.02) | (0.01) |
| Lockdown period (=1) | -0.31*** | -0.24*** |
| | (0.03) | (0.02) |
| Phase I re-opening (=1) | -0.08*** | -0.08*** |
| | (0.03) | (0.01) |
| Phase II re-opening (=1) | -0.04** | -0.01 |
| | (0.02) | (0.02) |
| Observations | 16,650 | 16,650 |
| R-squared | 0.98 | 0.95 |
| Neighborhood FE | Yes | Yes |
| Event-month FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table replicates the analysis in Table 2 but uses a one-month window instead of a two-week window to define commuters. The dependent variables in columns (1)-(2) are the number of commuters (in logarithm) and average working hours for commuters (in logarithm), respectively. An individual who visits his work location at least once in a given month is coded as a commuter, and an individual who does not visit his work location at all during that month is coded as a non-commuter. Both columns include neighborhood, event-month, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at event-month. $*$ p <0.1, $**$ p < 0.05, $***$ p < 0.01.

**Table A.6. Effects on number of individuals making unemployment calls and call duration (excluding people under 25 years old)**

| VARIABLES | (1)<br>No. of individuals making calls (in log) | (2)<br>Duration per call (in log) |
|---|---|---|
| **Panel A:** | | |
| 1-30 days before lockdown | 0.03 | 0.03 |
| | (0.03) | (0.04) |
| Lockdown period | -0.38*** | -0.35*** |
| | (0.06) | (0.07) |
| Phase I re-opening | 0.26*** | 0.22*** |
| | (0.02) | (0.05) |
| Phase II re-opening | 0.43*** | 0.58*** |
| | (0.02) | (0.05) |
| **Panel B:** | | |
| Pandemic period | 0.27*** | 0.32*** |
| (Lockdown + Phase I + Phase II) | (0.04) | (0.03) |
| Observations | 489,514 | 489,514 |
| R-squared | 0.82 | 0.59 |
| Neighborhood FE | Yes | Yes |
| Event-day FE | Yes | Yes |
| Day-of-week FE | Yes | Yes |
| Holidays FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table replicates the analysis in Table 1, except that individuals under 25 years old are excluded from the analysis. The dependent variables in columns (1)-(2) are the number of individuals making unemployment calls (in logarithm) and the average duration of unemployment calls (in seconds, in logarithm), respectively. Both columns include neighborhood, day-of-week, event-day, holiday, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at the event-day. $*$ p <0.1, $**$ p < 0.05, $***$ p < 0.01.

**Table A.7. Regression results on commuters and working hours (excluding people under 25 years old)**

| VARIABLES | (1)<br>No. of commuters (in log) | (2)<br>Working hours (in log) |
|---|---|---|
| 1-30 days before lockdown (=1) | 0.01 | 0.01 |
| | (0.02) | (0.01) |
| Lockdown period (=1) | -0.29*** | -0.21*** |
| | (0.03) | (0.02) |
| Phase I re-opening (=1) | -0.09*** | -0.09*** |
| | (0.03) | (0.02) |
| Phase II re-opening (=1) | -0.04*** | -0.02 |
| | (0.02) | (0.02) |
| Observations | 34,965 | 34,965 |
| R-squared | 0.96 | 0.88 |
| Neighborhood FE | Yes | Yes |
| Event-fortnight FE | Yes | Yes |
| Treatment group FE | Yes | Yes |

**Notes:** This table replicates the analysis in Table 2, except that individuals under 25 years old are excluded from the analysis. The dependent variables in columns (1)-(2) are the number of commuters (in logarithm) and average working hours for commuters (in logarithm), respectively. An individual who visits his work location at least once in a given fortnight is coded as a commuter, and an individual who does not visit his work location at all during that fortnight is coded as a non-commuter. Both columns include neighborhood, event-fortnight, and the treatment group fixed effects. Standard errors are reported in parentheses and clustered at event-fortnight. $*$ p <0.1, $**$ p < 0.05, $***$ p < 0.01. Results are similar using one week or one month to define commuters.