Decoding Job Search Behaviour Using Search Volume Data*

Wei-Fong Pan[1]

First draft: 1 Nov 2018

This draft: 19 Aug 2019

Abstract

This study develops a set of job search (JS) indices based on China's Internet search data that reflect the level of JS and overcome the problem of survey data unavailability. Based on these indices, we document that JS primarily occurs among males and individuals aged between 20 and 39, and further demonstrate its seasonality. Considering regional differences, the results suggest that JS behaviors are concentrated in the Eastern region, which is the most developed, but are lower in less developed regions. Moreover, weekends and holidays also have negative effects on JS in general. Lastly, we find that the cyclicality of JS behavior varies by region and the device on which JS is conducted. Although the aggregate JS index suggests that JS is countercyclical, we find certain regions' JS to be procyclical or acyclical. Job seekers' usage of the computer or mobile phone for JS activity also demonstrates different cyclicality. These results might help to reconcile the mixed evidence of JS cyclicality in the literature.

Keywords: Job search; Business cycle; Internet search; China

JEL Classification: J20; J60; J61; J63

* Thanks to Iourii Manovskii, Yanling Yang, Shengfeng Lu, Junwei Shi, Renjun Shen, Ting Li, and seminar participants at 9th CEC Wuhan Seminar for their invaluable comments. This paper is previously circulated under the title "Understanding The Pattern Of Job Search Behavior". All errors are mine.

1 Department of Economics, School of Politics, Economics and International Relations, University of Reading. Email:wei-fong.pan@pgr.reading.ac.uk.

1 Introduction

Search effort is an important determinant of employment as a higher number of job matches are formed when both recruiters and job seekers put in more effort to find a suitable employee and job, respectively. Recruiting efforts have been studied by many researchers. For example, Davis et al. (2013) emphasized the importance of recruiting intensity; they show that it can deliver a better-fitting empirical Beveridge curve and accounts for a large share of the cyclical pattern of aggregate hires.

However, research on the potential employee's job search (JS) effort is limited, especially in terms of JS cyclicality and behavior in emerging countries. It is important to understand the role of cyclicality in the JS effort. The standard matching model suggests that if the job-finding rate increases with search effort, a countercyclical search effort dampens and a procyclical search effort amplifies the volatility of unemployment and vacancies.[2] However, empirical evidence on the cyclicality of JS activity is mixed. DeLoach and Kurt (2013), Leyva (2018), and Mukoyama et al. (2018) analyzed data from the American Time Use Survey (ATUS) to conclude that JS activity is either acyclical or countercyclical, similar to the result of Shimer's (2004) study, whereas Gomme and Lkhagvasuren (2015) found that JS is procyclical.[3] It is clear that the cyclicality of JS is an unresolved issue.

[2] This is why Mukoyama et al. (2018) suggested that a higher unemployment rate would be observed if the JS were not countercyclical.

[3] Some other studies indirectly observe countercyclical JS efforts. For example, Faberman and Kudlyak (2019) found that employees tend to search less in areas with tight labor markets, which implies that the JS effort is countercyclical. However, the estimates of Chodorow-Reich and Karabarbounis (2016) implied that the search effort is acyclical or less countercyclical.

Studies on JS in emerging countries remain limited. Among these, Banerjee and Bucci (1995), Munshi (2003), and Nicodemo and García (2015) used survey data to investigate the relationship between social networks and individual JS behaviors. Zenou (2008) theoretically showed that a hiring subsidy policy has a clear positive effect on formal employment, and that reducing the unemployment benefit has an ambiguous effect on employment in developing countries. Even though these studies try to fill the gap, there remains an unexplored area in understanding JS in emerging countries.

This study aims to provide some evidence on the JS characteristics in China, which is the world's largest developing economy. One challenge faced when studying JS behavior in China is the absence of relevant data. To overcome this difficulty, this study measures JS behaviors by using search volume data, similar to Da et al. (2011) and Baker and Fradkin (2017). This approach enables us to measure JS when time use data is unavailable or problematic (e.g., is characterized by low frequency and biased samples). We use Baidu search data instead of Google because Baidu is the largest search engine in China. Using Baidu data not only helps avoid the look-ahead bias that might exist in Google Trends as Google normalizes raw search data,[4] but also provides data on the characteristics, including gender and age, of users who search for specific terms and of individuals who use computers versus mobile phones for search, thereby providing more insights on this issue. Regarding the search terms, although Baker and Fradkin (2017) suggest that using "jobs" adequately represents the US JS behavior, it may not be an effective term in China as Chinese job seekers do not directly search "jobs" when they are indeed looking for jobs.[5] Instead, we

[4] We will discuss this issue in Section 2, when we construct our JS index.

[5] In section 2, we will show why the term "jobs" is inappropriate for modelling JS in China.

propose to use the names of job portals as search terms for JS. This is because individuals' searches for job portals imply that they are searching for jobs or are at least interested in JS. Using search volumes of the name of job portals as a proxy for the level of JS has another advantage in that we can observe the JS of a particular group of people (Pan, 2019).

This study makes several contributions to the literature. First, similar to Baker and Fradkin (2017), this study creates a set of JS indices, enabling researchers to investigate JS behaviors at higher frequencies and across different geographical regions. This contribution complements existing studies and allows interested parties to examine the labor market when conventional survey data are unavailable or deficient. In particular, our index differentiates between the JS efforts of individuals who use computers and mobile phones.

Further, we comprehensively investigate JS efforts in China. First, similar to other labor market variables, JS behavior demonstrates seasonality; we observe spikes generally in March and the lowest levels in January of each year. Second, we also observe that JS is more frequent among male workers; however, the gap between male and female job seekers narrows. The group aged between 20 and 39 contributes 70% of internet JS activity. Third, JS behaviors are concentrated in the Eastern region, which is the most developed region, but is lower in less developed regions. Fourth, the pattern of JS differs when using personal computers versus mobile phones. Fifth, we conclude that JS behaviors are most intense on Mondays and lowest during holidays. Further, the search intensity decreases on Saturdays and increases insignificantly on Sundays.

Lastly, we investigate the cyclicality of the JS effort. We observe that cyclicality varies across regions and devices used for JS. Although we observe that aggregate JS tends to be countercyclical, which is similar to the case in the US uncovered by Mukoyama et al. (2018), the results are significantly different when considering regional variations. We observe that JS in the central region tends to be procyclical while that in the northeastern region is countercyclical. The device used for searching for jobs online also affects the cyclicality of the JS effort. Personal computer (PC)-based JS tends to be procyclical, but mobile-based JS tends to be countercyclical. These two findings explain that regional differences and devices might affect cyclicality and thereby, reconcile the mixed evidence found in the literature.

The rest of this paper proceeds as follows. In Section 2, we outline the method employed to construct the JS index and provide our data sources. In this section, we also test the robustness of our index. Section 3 studies the characteristics of JS behaviors using JSI and verifies its cyclicality. Section 4 concludes the paper and suggests several new lines of research.

2 Job Search Measurement

2.1 Previous Job Search Indicators

Two traditional indicators, unemployment and the employment-population ratio, are commonly used to measure JS activity. While each has its own advocates, both measures have drawbacks (Shimer, 2005). This is especially true of the employment-population ratio, which is a more cyclical measure of the JS activity, a fact that exacerbates Shimer’s critique.

Until recently, researchers studied JS behavior using time-use survey data. The most popular dataset is the ATUS from the Bureau of Labor Statistics. The intended use of this dataset is to examine how individuals allocate behaviors under different business cycles or financial conditions. Nevertheless, the method requires detailed survey data, which restricts the reach of relevant studies outside the US.

Mukoyama et al. (2018) overcame this problem by linking ATUS and Current Population Survey (CPS) data. Specifically, they used ATUS data to construct the imputed JS time for the respondents to the CPS survey. They further suggested computing the extensive and intensive margins to represent JS intensity. The extensive margin is the number of unemployed working-age individuals relative to total non-employment, while the intensive margin is the average time spent by the unemployed on JS activities. Mukoyama et al. (2018) offer a viable approach to developing an indicator based on labor market variables; nonetheless, studies in developing countries remain limited by data unavailability. Therefore, we cannot verify the time use behaviors of job seekers in countries without detailed labor market statistics.

2.2 Construction of the Job Search Index based on Internet Searches

Human attention is undoubtedly scarce because one is confronted with many options when making decisions (Kahneman, 1973; Barber and Odean, 2007). Da et al. (2011) consider search engine query data as a direct measure of individual attention and use data on stock ticker names from Google Trends as a proxy for investor attention to the stock market. They show that investor attention is an important asset-pricing factor. In Stephens-Davidowitz's new book, "*Everybody lies: Big data, new data, and what the internet can tell us about who we really are*" (Stephens-Davidowitz, 2017) and academic paper (Stephens-Davidowitz, 2014), the author also advocates the usage of Internet search data because survey respondents may arguably not disclose their true answers and even lie, which leads to biased survey results. This study is not the first one to attempt to use Internet search data to investigate the labor market. The seminal work of Baker and Fradkin (2017) uses the Internet search volume of job-related keywords in Google Trends as a proxy for aggregate JS behavior, to evaluate unemployment insurance policies for job seekers. Based on these reasons, we consider Internet search volume to be a good proxy for JS levels and use it in this study.

Since the study sample is from China, we use the Baidu Index from Baidu, the largest search engine in China.[6] Similar to Google Trends, the Baidu Index allows ordinary users to query the search frequency of a specific keyword on Baidu over different periods as well as its changing frequency trends.[7] Using Baidu search data has at least one advantage over survey-based data; it is based on millions of Internet users, which can be aggregated across geographies and at a higher frequency. Moreover, the Baidu Index has at least three advantages over Google Trends. First, it provides the actual frequencies of searches over time without normalizing them,[8] which can help avoid the look-ahead bias that might exist in Google Trends data. Second, it allows us to understand those who search for specific terms by providing information on user characteristics, including gender and age. Third, Baidu classifies the search volumes by device (i.e., PC-based and Mobile-based), which helps us to investigate whether JS using different devices may differ.

Regarding keyword search, one may first consider "jobs" in Chinese as the primary proxy for JS behavior. However, unlike in the US, this is not an appropriate word to measure JS behavior in China as job seekers do not directly search for "jobs" when they are looking

---

[6] See http://gs.statcounter.com/search-engine-market-share/all/china. One reason for Google's very low market share in China is the Chinese government's official ban on the use of Google.

[7] Although the Baidu Index does not have an English version, a user tutorial is available at https://sampi.co/baidu-index-tutorial/.

[8] Google Trends normalizes each search term to range from 0 to 100 for a given sample period, where 100 represents the date that the given search term achieves its peak relative search volume. Therefore, changing the sample period yields different values for the search volume indices of each term.

for a job. Figure 1 reports the Baidu Index demand-mapping function to confirm its inappropriateness. A demand-mapping function is used to reflect the actual demand of internet users who search for specific words ("jobs" in this case, "工作" or "职位" in Chinese).[9] The top panel of Figure 1 shows the Internet search demand before and after the user searches for "工作." Specifically, the left-hand side shows the keywords searched by Internet users before they search for "工作." Here, the top five searches are for "head class teacher," "Party members," "working report," "regulations," and "China." The right-hand side shows the keywords searched for by individuals after searching for "工作." Here, the top five keywords are similar to the left-hand side, namely "head class teacher," "regulations," "half-year," "party members," and "China." Next, the bottom panel of Figure 1 summarizes the Internet search demand before and after the user searches for "职位." It shows that people who search for "职位" are mainly interested in searching for jobs in the public sector. Hence, as evident, using the Chinese equivalent of "jobs" is inappropriate in our case to measure the level of JS in China.

Instead, we follow Pan (2019) who advocates using the names of job portals to measure the JS levels. This is because individuals' searches of job portals indicate that they are searching for or are at least interested in jobs. We select two job portals, namely Zhilian (智联招聘) and qianchenwuyou (前程无忧) as they are well-known and established much earlier than others. Lastly, we aggregate the search volumes ($SV_{i,t}$) of each term to create the JS index (JSI) by:

$$JSI_t = \sum_{i=1}^{N} SV_{i,t}. \quad (1)$$

[9] There are several alternatives for "jobs" in Chinese; hence, we check each of them successively.

As mentioned, Baidu differentiates the search frequencies based on the device: PC or mobile phone. Figure 2 plots the aggregate JS for both device types; the mobile-based JSI is from 1 January 2011 to September 2018 because of data availability. We plot monthly observations of the JSI with the business condition index (BCI), which serves as a proxy for business cycle fluctuations. The PC-based JSI has several spikes with the highest value recorded in February 2012, which corresponds to the concerns regarding China's slowdown in growth and decline in real economic activity. This observation can be verified by the high negative correlation between BCI and JSI (-50%). Thereafter, the PC-based JSI remained at a high level from 2013 to the beginning of 2017. In comparison, mobile-based JSI has a similar observation as it achieved its highest value in September 2015, which corresponds to the 2015 stock market crash and further slowdown of economic growth in China. However, it exhibits some differences with PC-based JSI; for example, there is no obvious spike in 2013 in mobile-based JSI. This divergence might arise from the fact that using mobile phones to search for jobs was not prevalent in 2013.

2.3 Validity of our Job Search Index

In this subsection and the next, we perform robustness checks for our index. For brevity, we provide only country-level JSI as evidence, but using province-level JSIs qualitatively leads to the same conclusion.

Extension of keywords

We notice that more job portals have been established in recent years including 51jobs.com and 58Tongchen. This raises the concern that using only two specific job portals as search terms may be biased. To address this concern, we include four additional job portals to create an alternative JSI. Panel A of Figure 3 compares the benchmark index with this alternative. It is clear that using an extended set of keywords does not significantly change the two benchmark indices. The benchmark PC-based JSI is highly correlated with its alternative at over 98%. Even if we consider the period from January 2013 to September 2018 (we do this subsample analysis because all the job portals used in the extension were established since 2013), the correlation is high at over 97%. A similar conclusion holds for mobile-based JSI. Hence, we conclude that the selection of the two oldest job portals is appropriate to reflect the level of JS.

Adding Google search volume

An additional concern is that the Baidu search engine might not have adequate market share, especially before 2010, as Google, the largest search engine in the world, still operated in China until 2010.[10] According to StatCounter statistics, the market share of the Google search engine exceeded 50% from 2006 to 2010, and was about 30% in 2012 and 10% in 2013. To address this concern, we include search volume data from Google Trends for the above search keywords and restructure a new index based on both Baidu and Google's search data. Panel B of Figure 3 compares the JSI including Google data with the benchmark JSI.[11] Clearly, including the Google search data does not significantly change the JSI; the overall pattern is very similar. The benchmark PC-based JS is highly correlated with the JS including

[10] In January 2010, Google announced that they were no longer willing to censor searches in China and would pull out of the country completely if necessary as shared in https://www.theguardian.com/technology/2010/jan/12/google-china-ends-censorship.

[11] We aggregate the indices using principal component analysis by following Pan (2019). See http://www.eviews.com/help/helpintro.html#page/content%2Fgroups-Principal_Components.html%23ww171280.

Google search data, with a correlation of around 85%. The correlation between the benchmark mobile-based JS, and its corresponding alternative is even higher at over 90%.

Reducing the effect of the long-run trend

The raw search frequencies might have long-run trends driven by style changes, popularity of JS websites, popularity of using mobile phones instead of computers to search for jobs, and so on, which are uncorrelated with JS behavior (Püttmann, 2018). These factors may lead to biases in constructing JS indices. To mitigate this concern, we de-trend all search frequencies and reconstruct the JSI. We primarily employed the de-trending method proposed by Hamilton (2018), which uses a cycle length of two years (i.e., 24 months) for monthly observations.[12] This method overcomes the problems of the Hodrick–Prescott filter, which produces a series with spurious dynamic relationships with no basis in the underlying data-generating process.[13] We reconstruct the JSI based on the de-trended series. Panel C compares the benchmark JSI with these alternatives, and it is clear that both benchmark and corresponding alternative indices have a similar pattern. Overall, based on the above validity checks, the benchmark construction is not an artifact of arbitrary choices made in its construction.

3 Characteristics of Job Search Activity

In Section 2, we discussed how our JSI captures JS activity and verified its robustness in

[12] The Hamilton filter involves conducting an ordinary least squares (OLS) regression of the variable at date $t + h$ on the four most recent values of date t to avoid these drawbacks and obtain a cyclical component series. The OLS regression is as follows: $x_t = \beta_0 + \beta_1 x_{t-h} + \beta_2 x_{t-h-1} + \beta_3 x_{t-h-2} + \beta_4 x_{t-h-3} + v_t$, where the cyclical components are the residuals, v_t.

[13] There is another reason to avoid using the HP filter in our case; the HP filter relies on the full sample period, which may introduce look-ahead bias in our index.

various ways. This section employs the JSI and search frequencies to examine JS behavior in China.

3.1 Male versus Female

It is natural to compare the JS levels by gender because of the existence of gender differences in wage and job mobility (e.g., Loprest, 1992). Moreover, some empirical studies reveal gender discrimination in job advertisements (Kuhn and Shen, 2012) in China and suggest that the Chinese society might prefer male over female employees (Bulte et al., 2011). Figure 4 compares the gender-wise percentage of individuals who search for JS-related words to total search frequencies. Before 2015, male Internet users contributed around 75% of total searches. However, the proportion of females that searched for jobs using the name of job portals as keywords has increased since 2016. The gap between the genders has narrowed since 2016, with males and females accounting for about 60% and 40% of the total searches, respectively. Overall, those findings are consistent with literature that females have different job mobility from males (e.g., Light and Ureta, 1992; Loprest, 1992; Keith and McWilliams, 1999) and with the traditional view of Chinese culture whereby male members should work while females must stay at home or find a stable job.

3.2 Job Search across Age Groups

Figure 5 compares the job search frequencies by age. We classify searchers into five groups: below 19, 20 to 29, 30 to 39, 40 to 49, and above 50.[14] The figure provides the proportional contribution of each group to total search volume during the period 2013-2018.

[14] We classify searchers into these five groups due to data availability.

Clearly, individuals aged 20 to 29 contributed the most, accounting for nearly 50% of the total searches from 2013 to 2015. The 30-39 age group contributed the second largest proportion at 40% before and during 2015. However, the proportion contributed by this age group increased from 41% in 2015 to 59% in 2018, surpassing the 20-29 age group (24% in 2018) in that year. The under 19 years and over 50 years age groups separately contributed less than 4% each, every year. These findings suggest that individuals aged 20-39 are the main online job seekers. This seems intuitive as people aged over 40 might not be familiar with Internet searches (compared to the 20-39 age group) and may not need to search for jobs using job portals given that they may be directly contacted by job hunters owing to their long work experience.

3.3 Seasonality

As can be observed, both plots in Figure 2 seem to demonstrate seasonality. To check whether job search exhibits seasonality, Table 1 reports average JS for each month. Taking a closer look, we observe increases in JS generally in March and the lowest levels in January of each year. This is intuitive because individuals may wait to receive their annual bonus at the fourth quarter before voluntarily quitting, and firms also adjust their number of employees during this time.[15] During June and August, the JS is also higher, which may result from the participation of fresh graduates in the labor market.

3.4 Regional Job Search Characteristics

Thus far, we have primarily focused on aggregate JS levels, but this might ignore the

[15] Many corporations in China distribute annual bonuses before or right after the Chinese New Year vacation, which usually falls in January or February of each year.

regional heterogeneity of JS behavior. As Blanchard and Katz (1992) and Manning and Petrongolo (2017) highlight the importance of regional heterogeneity in labor markets, this subsection presents the basic properties of regional JS behavior.

Figures 6 and 7 respectively display the local PC-based and mobile-based JSI. As shown in Figure 6, the pattern of JS varies across provinces.[16] For example, we observe a spike in 2012 in most provinces while others including Tibet and Guizhou do not have this spike. A similar pattern is observed in mobile-based JS indices. Most provinces' JS indices exhibit a sharp increase in JS in 2015, while Anhui province does not experience a spike in 2015 but in 2017.

Another approach to analyze the differences in JS patterns is to calculate their pairwise correlations. Table 2 shows the cross-province correlation plot for JSIs. The table shows considerable co-movement among JSIs. There are some exceptions, however. The PC-based JS in certain provinces, such as Guizhou, Jilin, Qinghai, and Liaoning, seem to have lower correlations with the rest of the economy. It is worth noting that the provinces that are observed to differ from the general pattern of PC-based JS are mainly areas that are less economically developed.[17] Next, mobile-based JS is typically highly correlated across provinces. Even in Anhui and Zhejiang, which exhibit among the lowest correlations of mobile-based JS, the correlation coefficients remain over 50%.

[16] Here we focus on 31 provincial-level divisions, excluding Hong Kong, Macau, and Taiwan as Baidu is their main search engine. These 31 provincial-level divisions consist of 22 provinces, four municipalities, and five autonomous regions. For brevity, we use the term "province" to refer to all the three different divisions involved in this study.

[17] Based on the National Bureau of Statistics of China in 2011, the Western and Northeastern areas are less economically developed (http://www.stats.gov.cn/ztjc/zthd/sjtjr/dejtjkfr/tjkp/201106/t20110613_71947.htm).

It is interesting to compare each province's PC-based and mobile-based JSI to assess whether there is an obvious divergence between these two types of JS behavior. As summarized in Table 3, it is clear that some provinces, for instance, Tibet, Chongqing and Guangxi, exhibit a high correlation coefficient of around 70%, while there is a weak positive correlation between PC- and mobile-based JS in many provinces with correlations ranging from 0% to 30%. Four provinces (Tianjin, Shanxi, Hebei, and Jilin) have weakly negative correlation coefficients. The weak correlation between the two types of JSI in many provinces seems intuitive because job seekers might choose to use only PC or mobile to search for jobs, and not necessarily both of them.

As mentioned in Section 2, given that Baidu provides the actual number of search frequencies, we can compare the level of JS for each province. Table 4 reports the descriptive statistics of JS for each province. A higher level of JS is concentrated in the Eastern region, which is the most economically developed area, while a lower level of JS is observed in less economically developed regions. For instance, Beijing city has the largest PC-based JSI at over 350,000 and Guangdong province has the largest mobile-based JSI at 358,319. Tibet has the smallest mean value of JSI for both devices. Such variations may arise from the population size and the proportion of the population with access to the Internet. To address this problem, we also compute the total searches per 1,000 Internet users.[18] Although JS per 1,000 users shows some different results from total JS, the general conclusion that JS is concentrated in more developed regions does not change. We continue to observe that the cities of Beijing, Shanghai, and Tianjin as well as Guangdong province have the highest JS per capita. Many other provinces in the Eastern region also have JS per capita levels that exceed the overall average. However, the average JS per user is lowest in the Western region, which is the least developed area in China. These findings are aligned with the intuition that

[18] Per capita JS indices are shown in the appendix. There is no pronounced difference between them and the corresponding original JS indices.

individuals have more employment opportunities in a region that is more economically developed. Overall, the results in this subsection emphasize the geographical differences in JS behaviors, which can be extended to understand the economic mechanism behind this phenomenon.

3.5 Weekday and Holiday Effects

Baker and Fradkin (2017) investigated weekday and holiday effects on JS efforts and found that JS activity drops significantly during weekends and is significantly lower on Fridays than on other weekdays. Since our JSIs can also be constructed using daily data, we perform the same analysis for China,[19] focusing on the effects of weekdays, weekends, and holidays.[20]

Table 5 shows the results of regressing provincial JSIs on the indicators for holidays and days of the week. Columns (1)–(2) show that PC-based JS efforts peak on Mondays and gradually decrease through the week. Similar to the US, Friday exhibits a lower level of search than other weekdays, but higher than that on weekends. PC-based JS efforts drop significantly during holidays. Weekends also have negative effects on its JS in general, but the effects are different on Saturdays and Sundays; Saturday has a positive, and Sunday has a negative effect on PC-based JS efforts. This finding suggests that Chinese job seekers using PC rest mainly on Sunday. We also observe similar findings when using mobile-based JS indices. Our findings are different from the US result of Baker and Fradkin (2017), where US JS efforts consistently decrease on weekends. These results are consistent when controlling

[19] We report the aggregate-level and province-level JSI using daily observations in the Appendix.

[20] Holidays include Chinese New Year, Qingming Festival, Dragon Boat Festival, Mid-Autumn Festival, and New Year’s Day.

for province and week fixed effects.

3.6 Is Job Search in China Countercyclical?

This section focuses on the cyclicality of JS behavior in China. As noted in the Introduction, the basic rationale for performing this analysis is that a procyclical JS effort amplifies the volatility of unemployment and vacancies, while a countercyclical one dampens it. Therefore, it is paramount to understand the cyclicality of JS behaviors, especially since the standard matching model predicts JS to be procyclical, while recent evidence suggests otherwise.

The basic approach to examining the cyclicality of JS activity is to investigate the time series property of the JSI and business cycle variables. As shown in Figure 2 in Section 2, there is a strong and negative correlation between aggregate JSI and BCI, which indicates that aggregate JSI (regardless of type) tends to be countercyclical.

However, does this conclusion hold true after considering regional heterogeneity? For robustness, we follow Aguiar et al. (2013), Haltiwanger et al. (2018), and Mukoyama et al. (2018) to regress the cyclical components of the JS index on the cyclical components of business cycle variables; this can be expressed as follows:

$$JSI_{it}^{c} = \mu_i + \mu_t + \beta_1 Cycle_{it} + \varepsilon_{it} \tag{2}$$

where JSI_{it}^{c} is the cyclical component of JSI, μ_i is the cross-sectional fixed effect, and μ_t is the time fixed effect. $Cycle_{it}$ is the business cycle measure for province *i* at time *t*. The business cycle variables are the filtered province-level GDP.[21] Moreover, the filtering

[21] One could argue that there are other alternatives such as unemployment, which can serve as proxy for regional business cycles. Here, we do not use the unemployment rate because it might be

approach is the same as in Section 2, but with different parameters. The province-level GDP is available at a quarterly frequency. As the Hamilton filter is based on a two-year sample length, we set h=8 for quarterly observations. β_1 is the coefficient of interest, where the cross-sectional fixed effect is used to capture any difference in JS behaviors across provinces.

Panel A of Table 6 reports the case for PC-based JSI while Panel B reports the results for mobile-based JSIs. The regression result suggests that there is a significant positive relationship between GDP and PC-based JSI using the full sample, indicating that PC-based JS is procyclical. The procyclicality of JS is observed at most regions, excluding the northeastern region. We observe the significantly negative relationship between GDP and JSI in the northeastern region, showing that its PC-based JS is countercyclical.

However, the results of the mobile-based JS are obviously different. The results suggest that mobile-based JS is weakly countercyclical as the coefficient β_1 is negative across most columns, but only significant in the case of the western region. In fact, β_1 remains significantly positive at the 5% level in the central region only.

The results summarized in Table 6 are quite different when comparing aggregate level BCI and JSIs; these offer at least two possible explanations of the conflicting evidence of the cyclicality of JS in the literature. First, regional differences might be a contributor, given that

problematic (Giles et al., 2005). The official unemployment rates are registered statistics and are not based on representative sample surveys. Among others, these figures ignore workers laid off with no expectation of reemployment and those who lost their jobs but did not register with their local governments (Solinger, 2001). We also compare the unemployment rate with the number of initial unemployment claims. Surprisingly, the unemployment rate is negatively correlated with number of initial unemployment claims. This comparison again shows the problematic nature of the official unemployment rate.

the cyclicality of JS varies across regions. If the survey data used by researchers contains a larger proportion of a certain area where the JS cyclicality differs from most other regions, then this might bias the results. This explanation is not only valid in China but also in other large economies, such as the US and Euro Zone, which consist of many disaggregate economic units. Second, most survey data related to JS do not appear to report the manner in which the job search is conducted (i.e., via PC, mobile phone, newspaper, or others). Our results show that online JS using different devices exhibit contrasting evidence of the cyclicality in JS.

4 Discussion and Conclusions

This study develops a new set of JS activity indices at the aggregate and regional levels based on internet search volumes and shows that it can capture JS activity in China. JS behavior is more intense in regions with better economic development. Although JS behaviors are highly correlated across provinces in general, it still differs across geographic areas to a certain extent. We also observe that males and individuals aged 20-39 are the main contributors of online JS activity. We also observe that JS behaviors are most intense on Mondays and lowest during holidays. Further, the search activity decreases on Saturdays and is resumed on Sundays.

After documenting several characteristics of JS behaviors, we study the cyclical properties of JS behavior using these indices. Although we observe that JSI tends to be countercyclical using aggregate-level JSI, we obtain very distinct results when considering regional differences. PC-based JS tend to be procyclical while mobile-based JS tends to be countercyclical. Moreover, JS varies across geographical regions. For example, the JS effort of job seekers using PCs in the northwestern region tends to be countercyclical, which is in contrast to the other regions.

We note mixed findings on the cyclicality of the JS effort; the results obtained offer at least two explanations for this mixed finding. First, regional differences may be a contributing reason as we observe that the cyclicality of JS varies across regions. Second, the device used by job seekers for job hunting could affect the JS effort. These imply that if the survey data used in existing studies contain larger proportions of job seekers from certain regions or a preference for using specific devices, they might yield a mixed portrait of JS behavior. These findings might help resolve Shimer's puzzle with respect to the fact that search intensity is not procyclical. Future research might use our approach to construct national or regional level JSIs for monthly or daily frequencies. Another possible line of research would be the determination of regional differences in the labor market. One of the limitations of our study is that this strategy cannot distinguish the JS efforts between the unemployed and on-the-job search. Future studies could explore this distinction.

Appendix

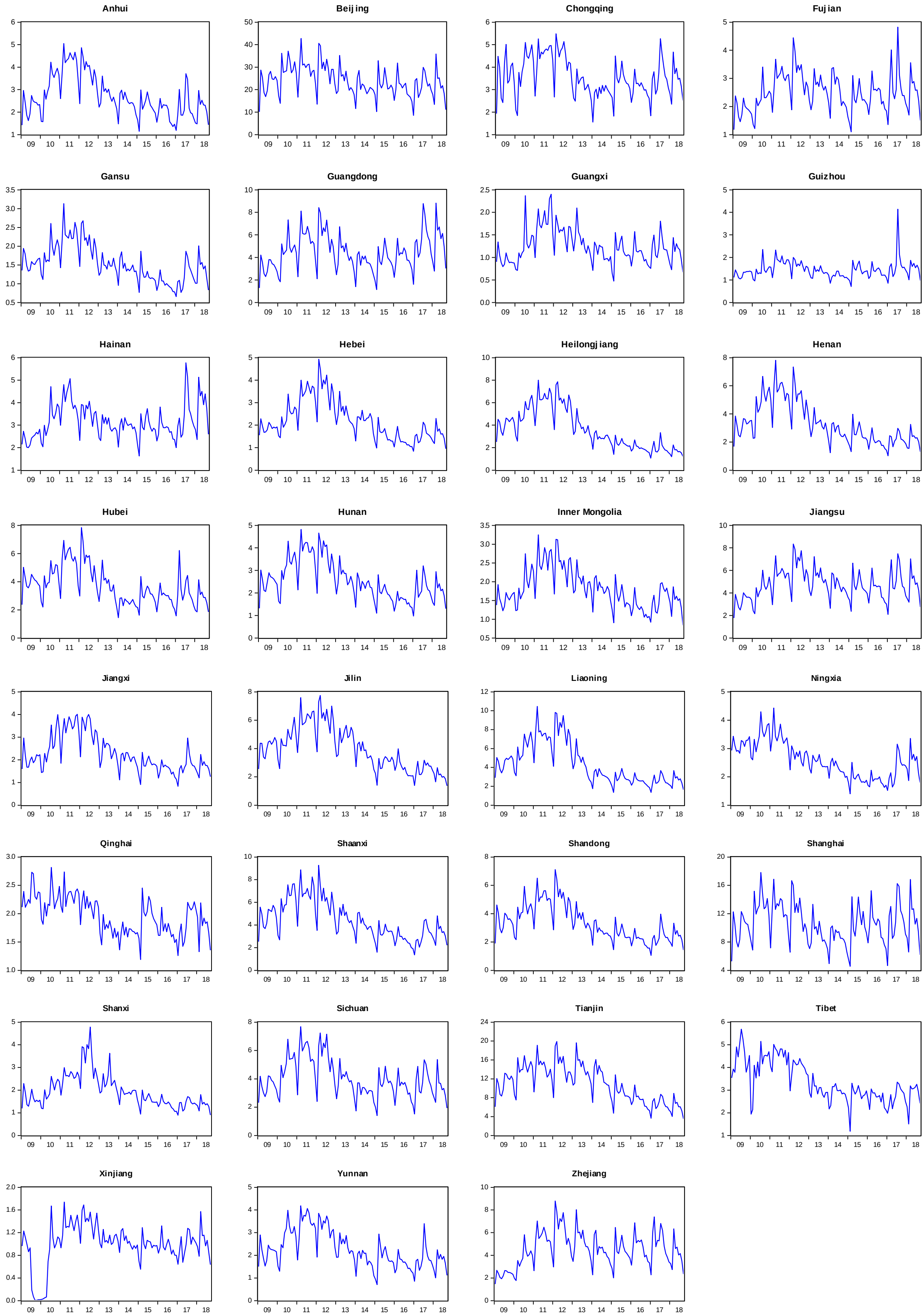


Figure A1. PC JS per capita

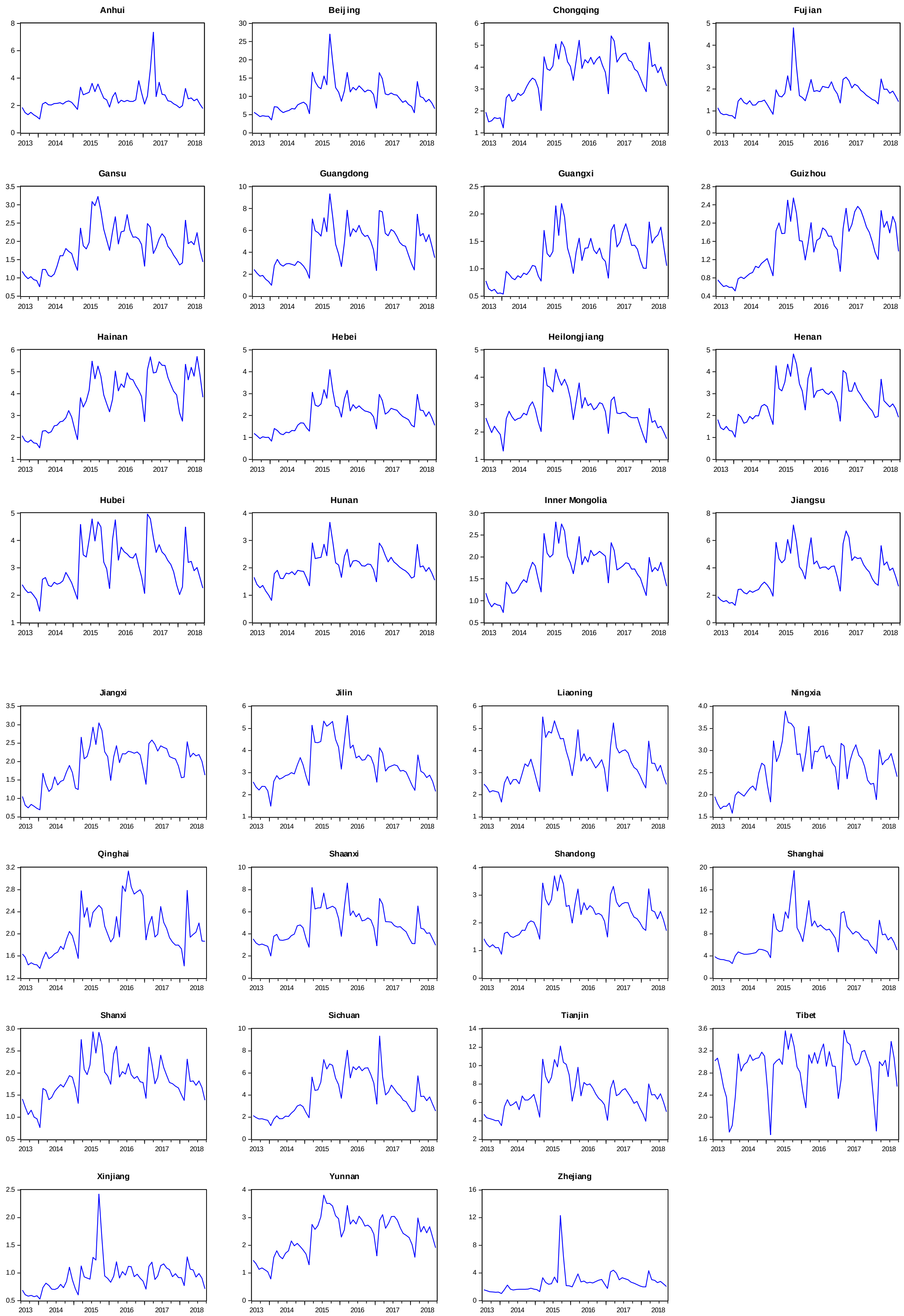


Figure A2. Mobile JS per capita

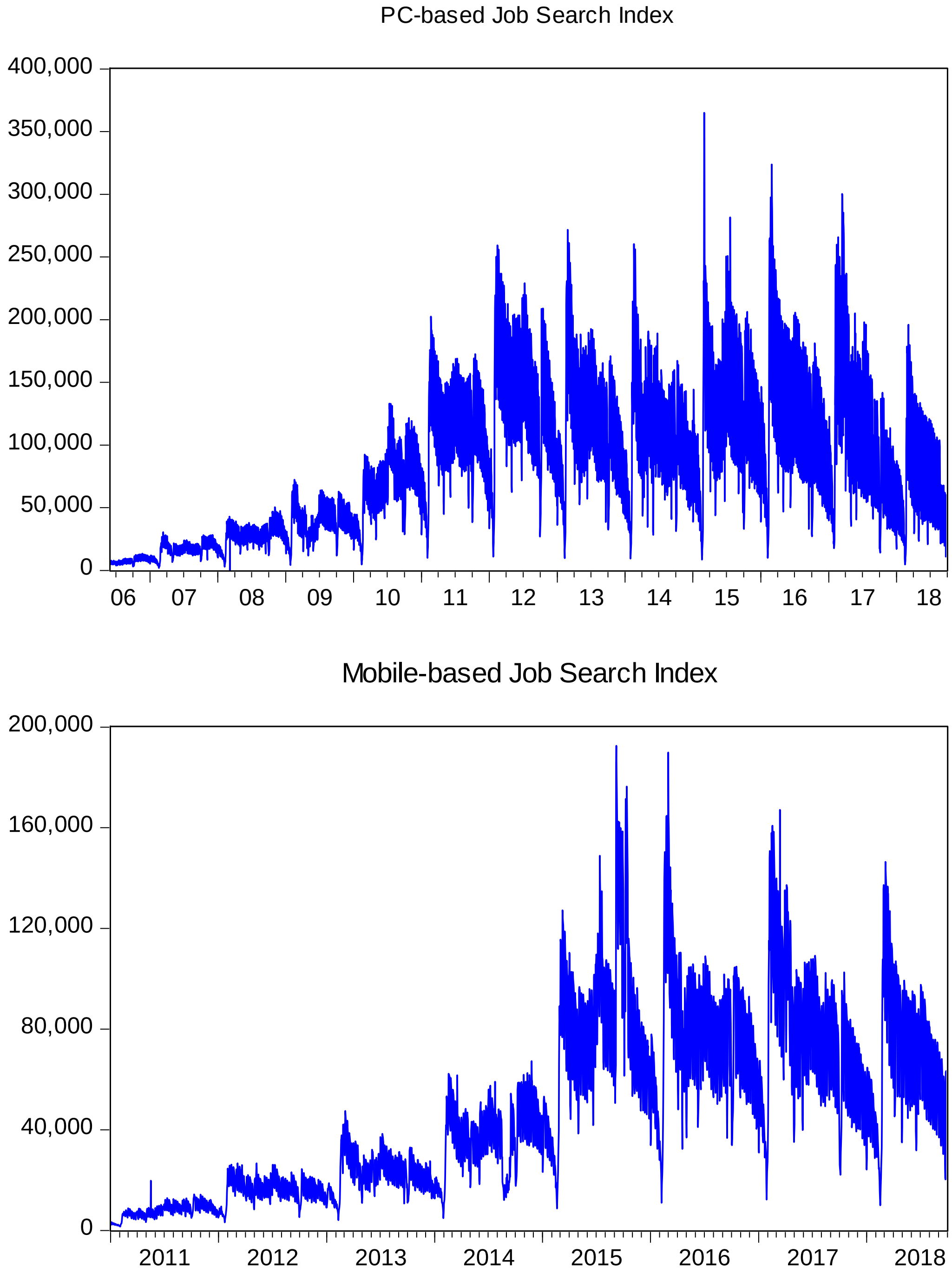


Figure B1. Daily Aggregate Job Search Indices

Notes: Top figure shows the PC-based job search index, while bottom figure shows the mobile-based job search index. Due to data availability, the sample period for top panel is from 1 June, 2006 to 30 September, 2019, and that for bottom panel is from 1 January, 2011 to 30 September, 2018.

Sources: Author’s calculations; Baidu

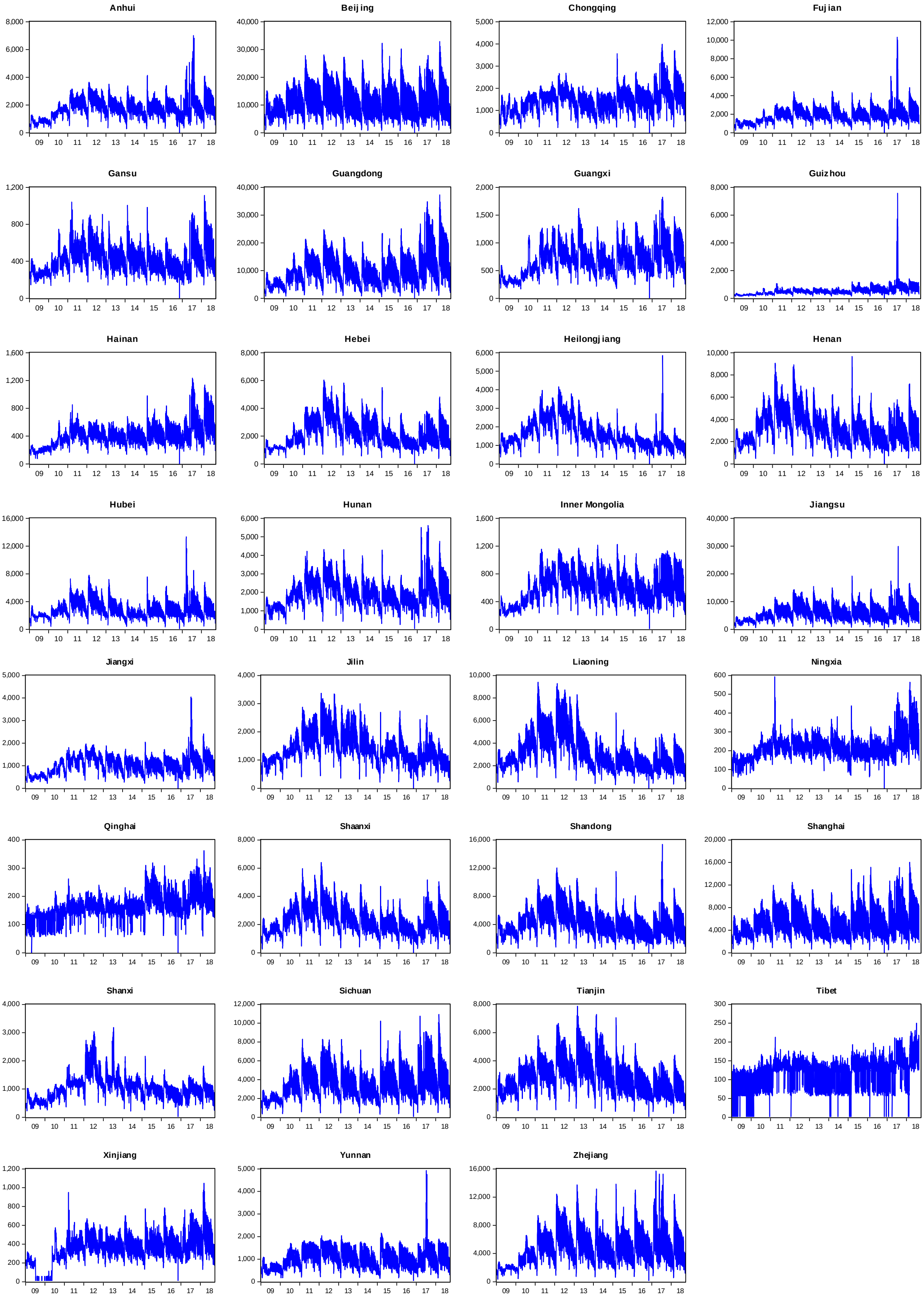


Figure B2. Daily PC-based job search index across province

Notes: The figure provides the PC-based job search for each province. The sample period begins from 1 Jan., 2009 to 30 Sep., 2018.

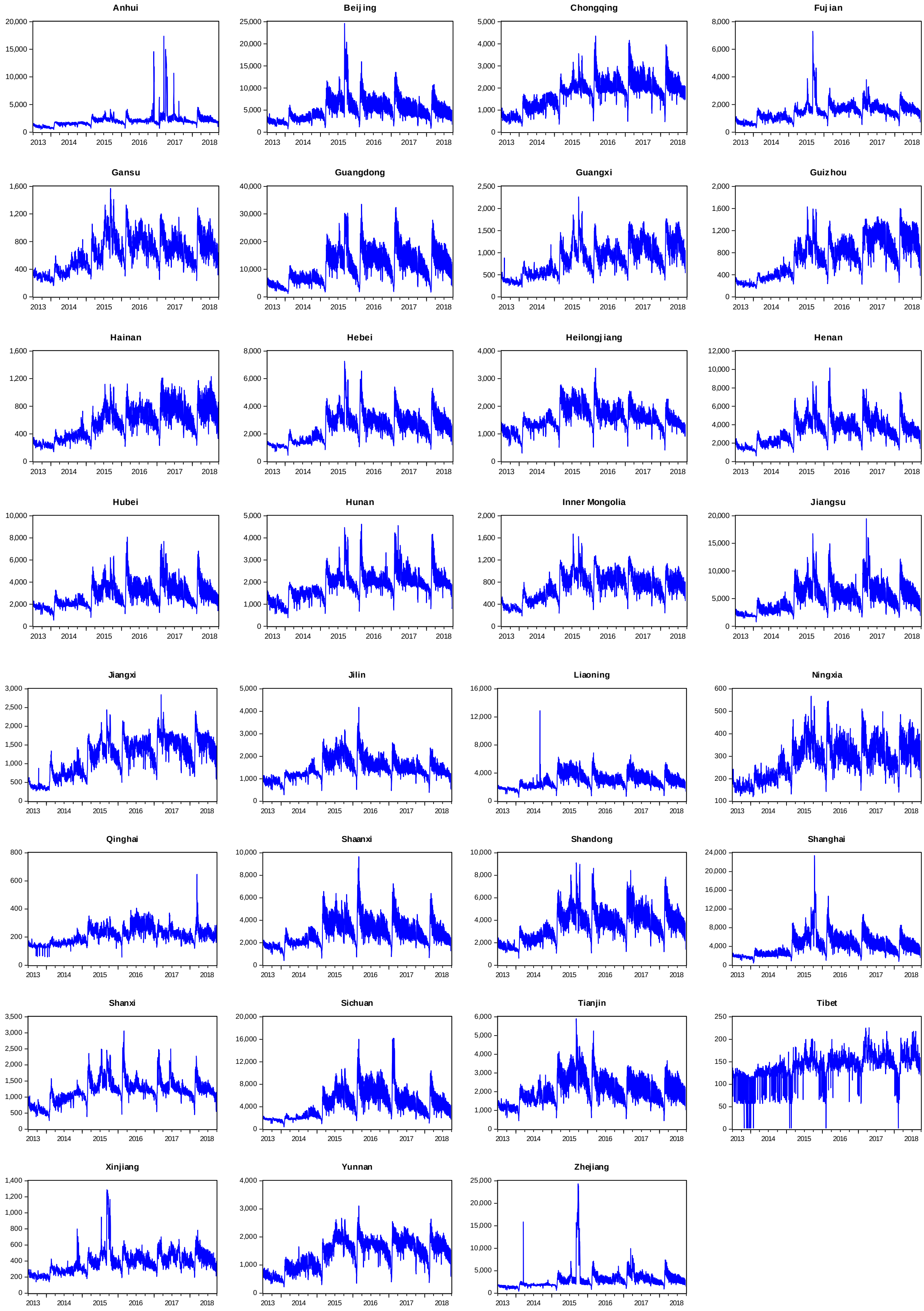


Figure B3. Daily mobile-based job search index across province

Notes: The figure provides the PC-based job search for each province. The sample period begins from 1 July 2009 to 30 September 2018.

References

Acemoglu, D. (2001). Good jobs versus bad jobs. *Journal of Labor Economics*, 19(1), 1-21.

Aguiar, M., Hurst, E., & Karabarbounis, L. (2013). Time use during the great recession. *American Economic Review, 103*(5), 1664-96.

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly, 55*(2), 107-120.

Baker, S. R., & Fradkin, A. (2017). The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data. *Review of Economics and Statistics*, 99(5), 756-768.

Banerjee, B., & Bucci, G. A. (1995). On-the-job search in a developing country: an analysis based on Indian data on migrants. *Economic Development and Cultural Change*, *43*(3), 565-583.

Bils, M., Chang, Y., & Kim, S. B. (2012). Comparative advantage and unemployment. *Journal of Monetary Economics*, *59*(2), 150-165.

Blanchard, O., & Gali, J. (2010). Labor markets and monetary policy: A New Keynesian model with unemployment. *American Economic Journal: Macroeconomics*, *2*(2), 1-30.

Bulte, E., Heerink, N., & Zhang, X. (2011). China's one-child policy and 'the mystery of missing women': ethnic minorities and male-biased sex ratios. *Oxford Bulletin of Economics and Statistics*, *73*(1), 21-39.

Chauvet, M., Gabriel, S., & Lutz, C. (2016). Mortgage default risk: New evidence from internet search queries. *Journal of Urban Economics*, 96, 91-111.

Chodorow-Reich, G., & Karabarbounis, L. (2016). The cyclicality of the opportunity cost of employment. *Journal of Political Economy*, *124*(6), 1563-1618.

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88, 2-9.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461-1499.

Davis, S. J., Faberman, R. J., & Haltiwanger, J. C. (2013). The establishment-level behavior of vacancies and hiring. *The Quarterly Journal of Economics*, *128*(2), 581-622.

DeLoach, S. B., & Kurt, M. (2013). Discouraging workers: Estimating the impacts of macroeconomic shocks on the search intensity of the unemployed. *Journal of Labor Research*, 34(4), 433-454.

D'Amuri, F., & Marcucci, J. (2010). 'Google it!'Forecasting the US unemployment rate with a Google job search index. FEEM Working Paper No. 31.2010. Available at SSRN: https://ssrn.com/abstract=1594132.

Faberman, R., & Kudlyak, M. (2014). The intensity of job search and search duration. Federal Reserve Bank of San Francisco Working Paper 2016-13.

Fondeur, Y., & Karamé, F. (2013). Can Google data help predict French youth unemployment?. *Economic Modelling*, *30*, 117-125.

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

Gomme, P., & Lkhagvasuren, D. (2015). Worker search effort as an amplification mechanism. *Journal of Monetary Economics*, 75, 106-122.

Haltiwanger, J. C., Hyatt, H. R., Kahn, L. B., & McEntarfer, E. (2018). Cyclical job ladders by firm size and firm wage. *American Economic Journal: Macroeconomics*, 10(2), 52-85.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics*, *100*(5), 831-843.

Kahn, L. M. (2012). Temporary jobs and job search effort in Europe. *Labour Economics*, *19*(1), 113-128.

Kahneman, D. (1973). Attention and Effort. Prentice-Hall, Englewood Cliffs, NJ.

Keith, K., & McWilliams, A. (1999). The returns to mobility and job search by gender. *ILR Review*, *52*(3), 460-477.

Krueger, A. B., & Mueller, A. (2010). Job search and unemployment insurance: New evidence from time use data. *Journal of Public Economics*, 94(3-4), 298-307.

Kuhn, P., & Shen, K. (2012). Gender discrimination in job ads: Evidence from china. *The Quarterly Journal of Economics*, *128*(1), 287-336.

Leyva, G. (2017). Against all odds: Job search during the Great Recession. Unpublished Manuscript.

Light, A., & Ureta, M. (1992). Panel estimates of male and female job turnover behavior: can female nonquitters be identified?. *Journal of Labor Economics*, 10(2), 156-181.

López-Salido, D., Stein, J. C., & Zakrajšek, E. (2017). Credit-market sentiment and the business cycle. *The Quarterly Journal of Economics*, *132*(3), 1373-1426.

Loprest, P. J. (1992). Gender differences in wage growth and job mobility. *The American Economic Review*, *82*(2), 526-532.

Manning, A., & Petrongolo, B. (2017). How local are labor markets? Evidence from a spatial job search model. *American Economic Review*, *107*(10), 2877-2907.

Meitzen, M. E. (1986). Differences in male and female job-quitting behavior. *Journal of Labor Economics*, 4(2), 151-167.

Mukoyama, T., Patterson, C., & Şahin, A. (2018). Job search behavior over the business cycle. *American Economic Journal: Macroeconomics*, 10(1), 190-215.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, *118*(2), 549-599.

Nicodemo, C., & García, G. A. (2015). Job Search Channels, Neighborhood Effects, and Wages Inequality in Developing Countries: The Colombian Case. *The Developing Economies*, *53*(2), 75-99.

Pan, W. F. (2019). Building sectoral job search indices for the United States. *Economics Letters*, *180*, 89-93.

Pissarides, C. A. (2009). The unemployment volatility puzzle: Is wage stickiness the answer?. *Econometrica*, *77*(5), 1339-1369.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, 118, 26-40.

Stephens-Davidowitz, S. (2017). Everybody lies: Big data, new data, and what the internet can tell us about who I really are. New York: HarperCollins.

Shimer, R. (2004). Search intensity. mimeo, University of Chicago.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1), 25-49.

Solinger, D. J. (2001). Why we cannot count the "unemployed". *The China Quarterly*, *167*, 671-688.

Zenou, Y. (2008). Job search and mobility in developing countries. Theory and policy implications. *Journal of Development Economics*, *86*(2), 336-355.

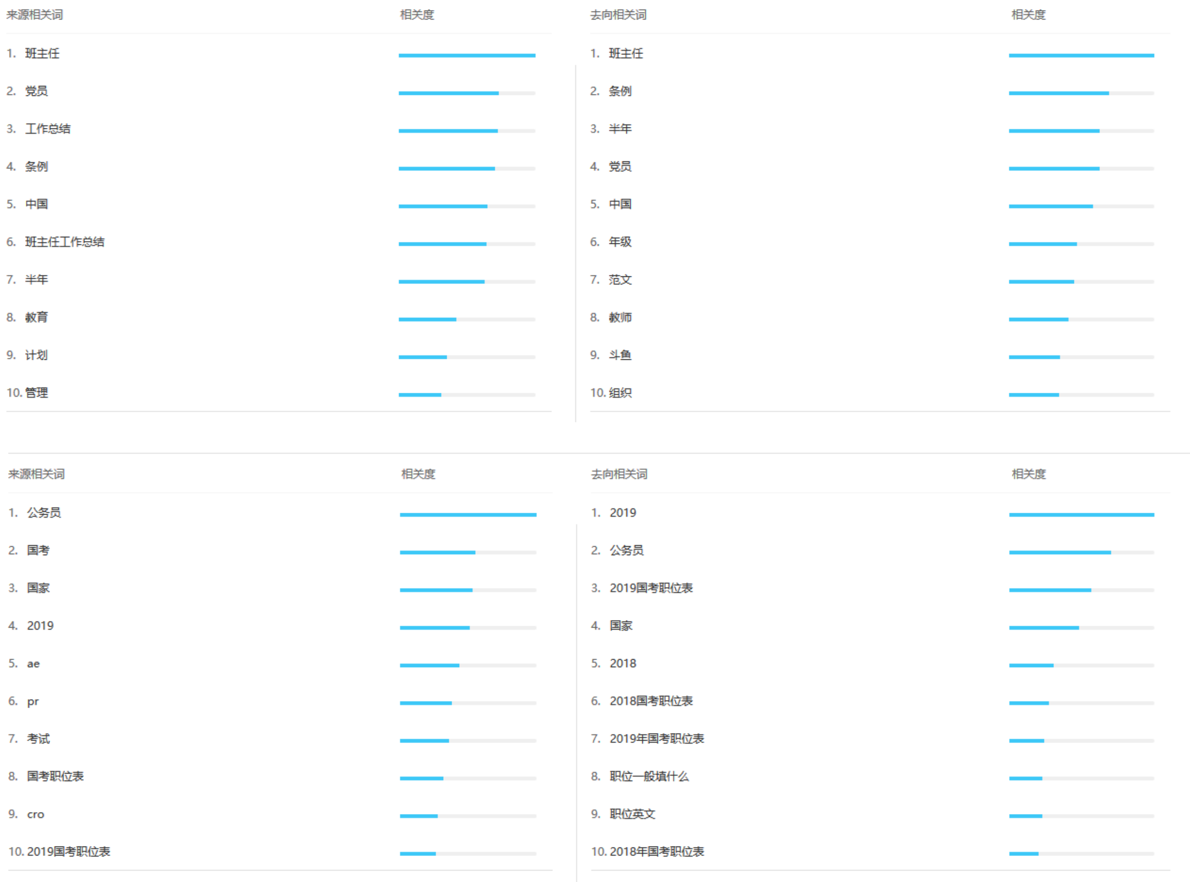
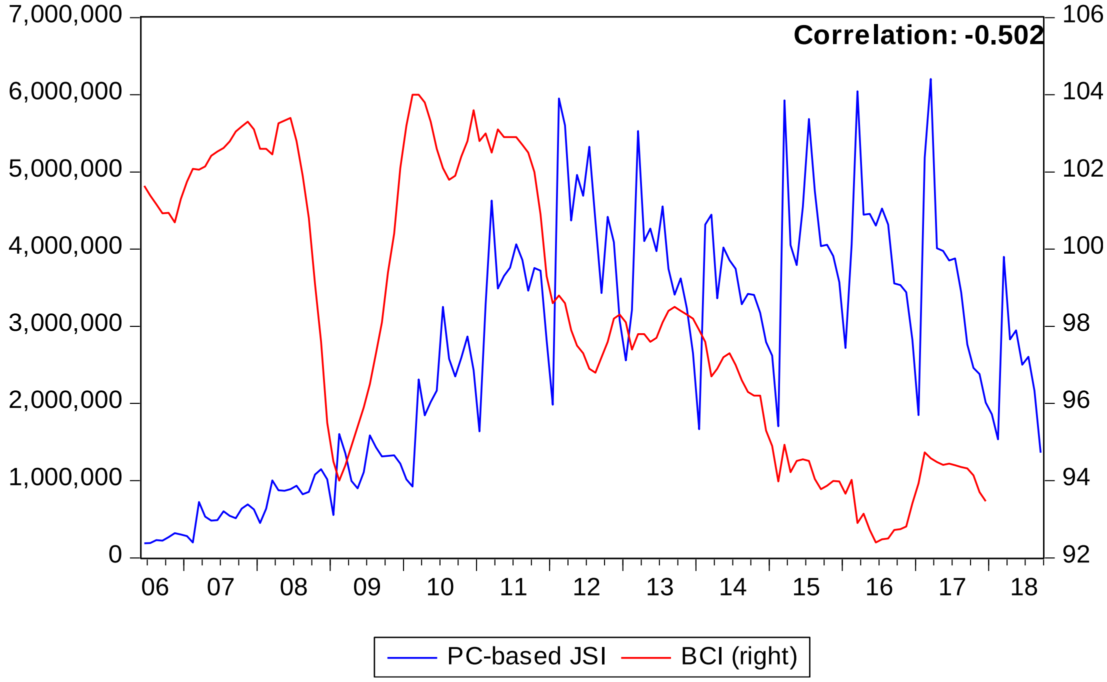


Figure 1. Illustrations of the demand mapping of "Jobs"

Notes: The top panel shows the keywords searched by Internet users before or after they search for 'gongzuo (工作)'. The bottom panel shows the keywords searched by Internet users before or after they search for 'zhiwei (职位)'. These two words are common alternatives for "jobs" in Chinese.



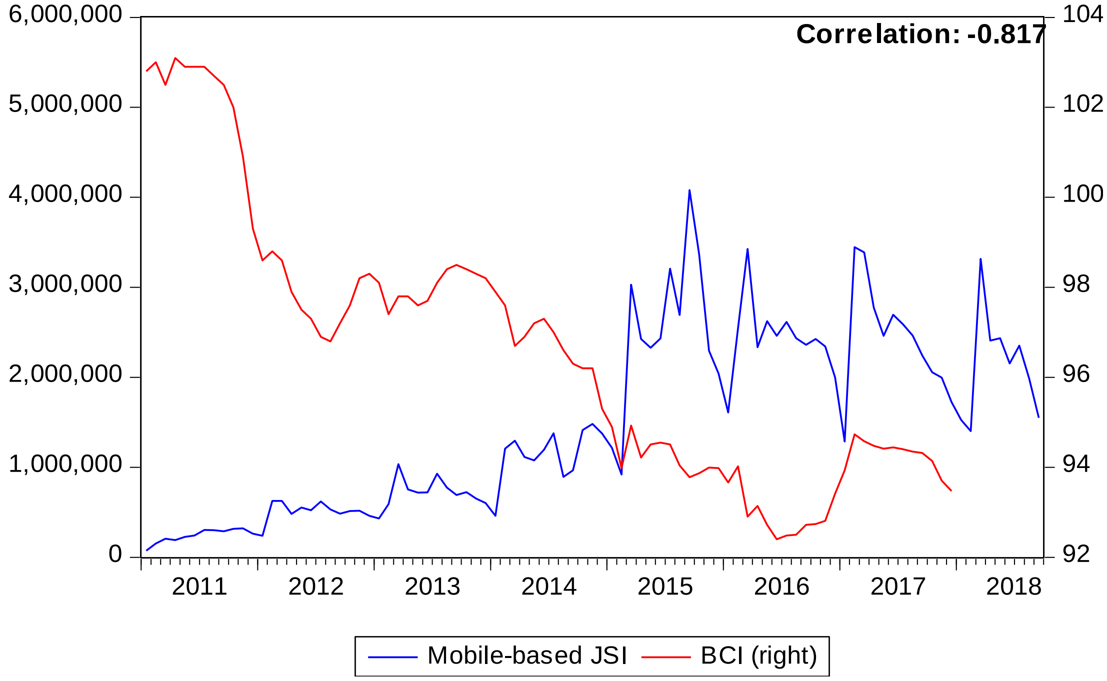


Figure 2. Aggregate Job Search Indices

Notes: Top figure shows the PC-based job search index, while bottom figure shows the mobile-based job search index. Due to data availability, the sample period for top panel is from June 2006 to September 2019, and that for bottom panel is from January 2011 to September 2018. BCI refers to business condition index from National Bureau of Statistics of China.

Sources: Author's calculations; Baidu; National Bureau of Statistics of China

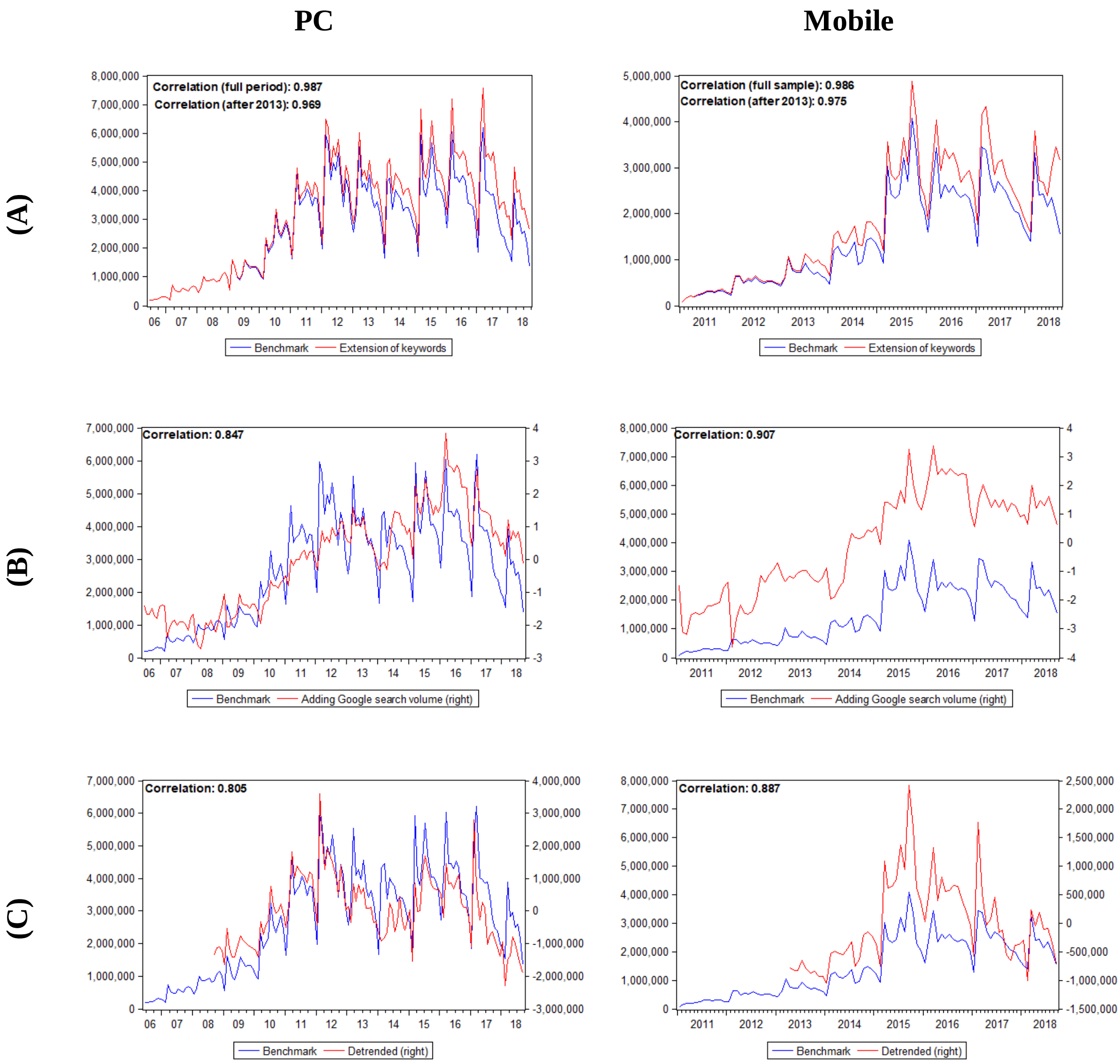


Figure 3. Comparison with other alternative JSIs

Notes: This figure compares benchmark index with index using alternative specifications. (A) extension of keywords; (B) Adding Google search volume; (C) using de-trended series.

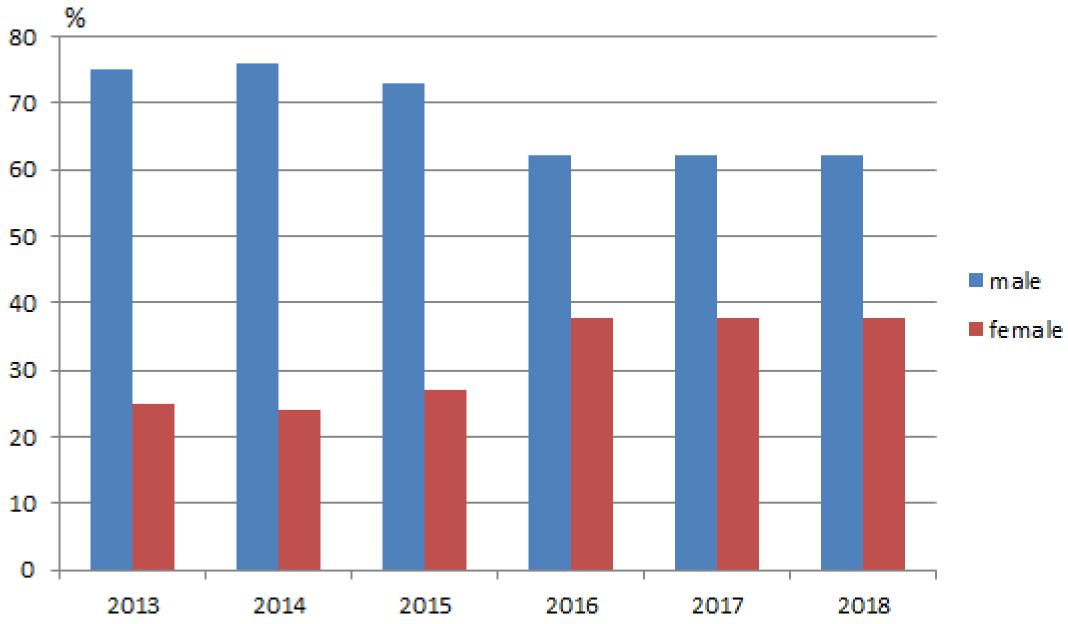


Figure 4. Job search behaviours in both males and females

Notes: The figure provides the proportional contribution of both gender to total search volume during the period of 2013-2018.

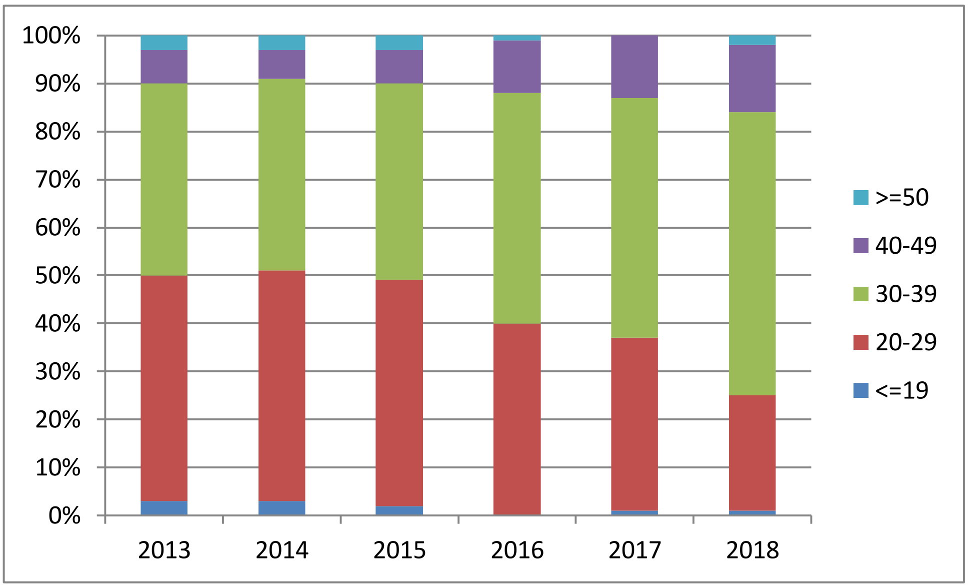


Figure 5. Job search behaviours in various age groups

Notes: The figure provides the proportional contribution of each age group to total search volume during the period of 2013-2018

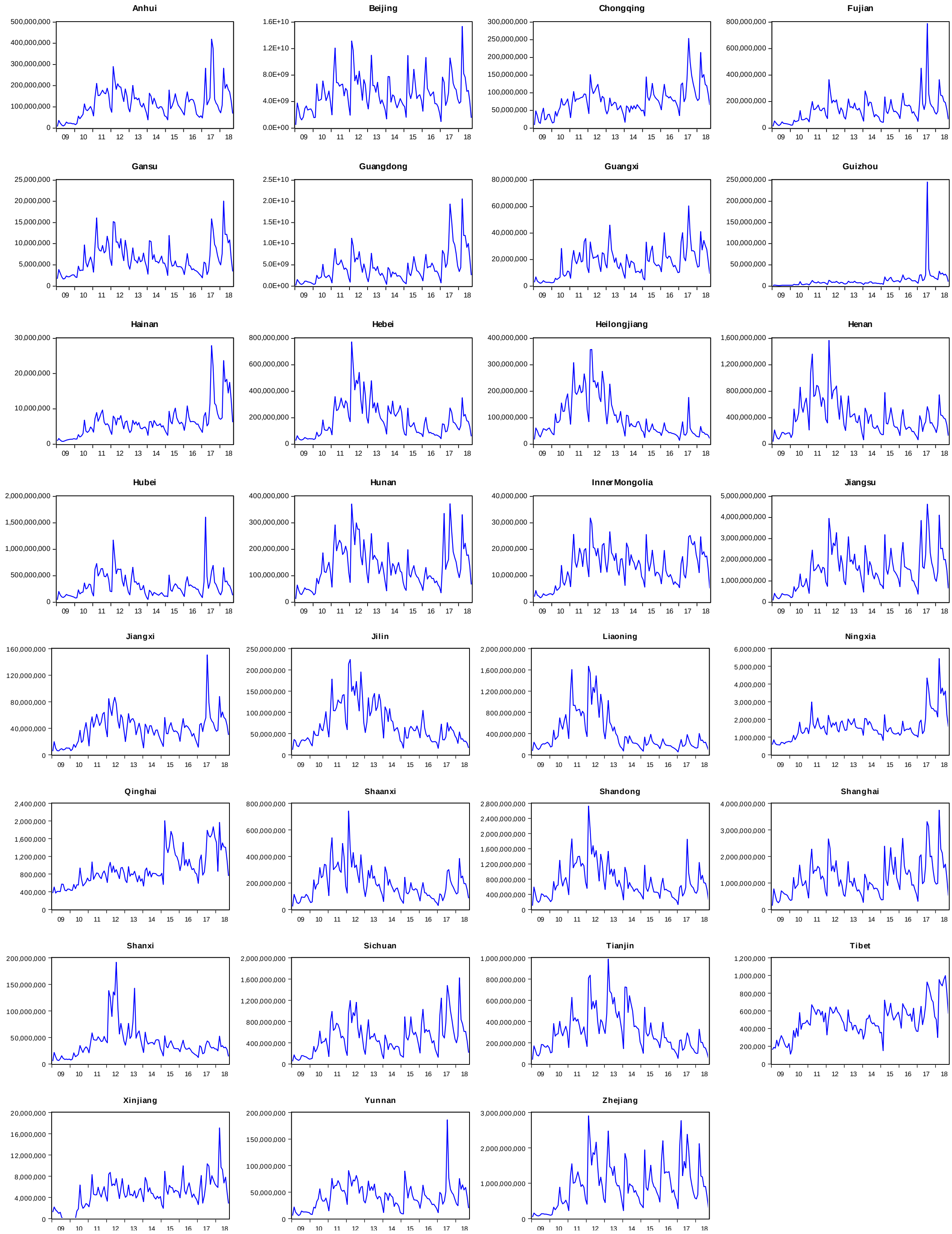


Figure 6. PC-based job search index across province

Notes: The figure provides the PC-based job search for each province. The sample period begins from Jan 2009 to Sep 2018.

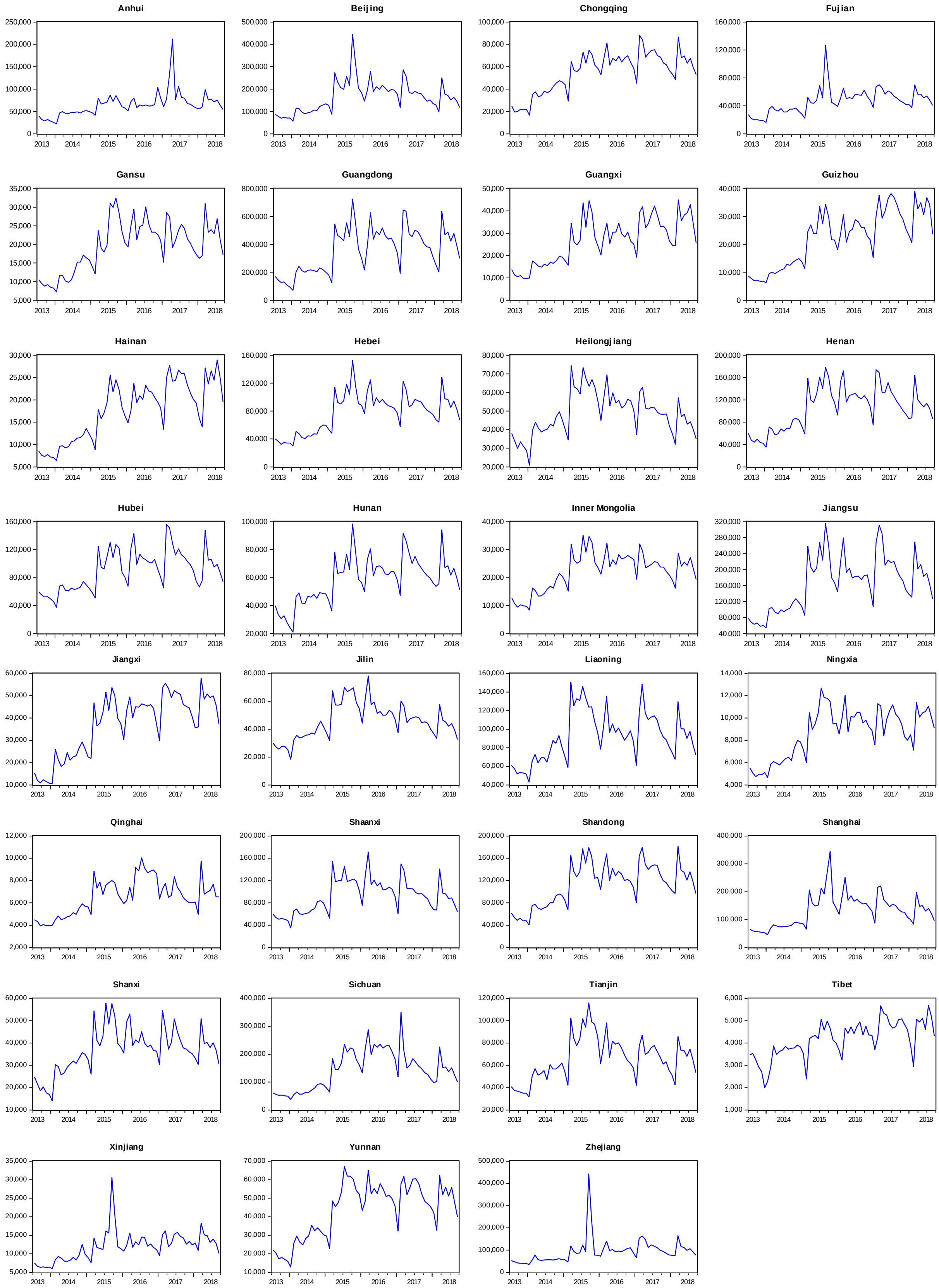


Figure 7. Mobile-based job search index across province

Notes: The figure provides the mobile-based job search for each province. The sample period begins from Jul 2013 to Sep 2018.

Table 1. The average JS index for each month

Month	PC	Mobile
January	14.082	13.275
February	14.476	13.782
March	14.986	14.191
April	14.692	13.949
May	14.712	13.974
June	14.524	13.982
July	14.652	14.129
August	14.539	13.973
September	14.410	13.957
October	14.518	13.962
November	14.523	13.893
December	14.367	13.765

Table 2. Correlations of job search index across province

	ANHUI	BEIJING	CHONGQING	FUJIAN	GANSU	GUANGDONG	GUANGXI	GUIZHOU	HAINAN	HEBEI	HEILONGJIANG	HENAN	HUBEI	HUNAN	INNER_MONGOLIA	JIANGSU	JIANGXI	JILIN	LIAONING	NINGXIA	QINGHAI	SHAANXI	SHANDONG	SHANGHAI	SHANXI	SICHUAN	TIANJIN	TIBET	XINJIANG	YUNNAN	ZHEJIANG
ANHUI	1.000	0.830	0.865	0.891	0.854	0.863	0.882	0.582	0.810	0.713	0.536	0.675	0.803	0.921	0.843	0.942	0.929	0.564	0.526	0.774	0.598	0.696	0.768	0.853	0.568	0.899	0.507	0.750	0.817	0.894	0.866
BEIJING	0.830	1.000	0.783	0.746	0.861	0.807	0.795	0.363	0.662	0.718	0.584	0.803	0.748	0.845	0.832	0.884	0.787	0.603	0.616	0.694	0.597	0.771	0.820	0.903	0.584	0.890	0.639	0.648	0.785	0.812	0.846
CHONGQING	0.865	0.783	1.000	0.834	0.774	0.939	0.857	0.657	0.903	0.472	0.273	0.502	0.649	0.771	0.727	0.865	0.873	0.284	0.292	0.827	0.808	0.507	0.561	0.916	0.389	0.907	0.203	0.847	0.842	0.827	0.750
FUJIAN	0.891	0.746	0.834	1.000	0.714	0.832	0.864	0.773	0.791	0.535	0.325	0.478	0.689	0.796	0.727	0.905	0.863	0.391	0.291	0.729	0.624	0.481	0.608	0.807	0.429	0.833	0.408	0.666	0.781	0.856	0.840
GANSU	0.854	0.861	0.774	0.714	1.000	0.819	0.746	0.406	0.740	0.774	0.623	0.777	0.657	0.892	0.868	0.821	0.843	0.621	0.618	0.820	0.534	0.823	0.827	0.786	0.585	0.802	0.553	0.692	0.810	0.813	0.696
GUANGDONG	0.863	0.807	0.939	0.832	0.819	1.000	0.839	0.634	0.929	0.503	0.271	0.515	0.646	0.795	0.728	0.862	0.843	0.284	0.315	0.899	0.728	0.529	0.585	0.917	0.378	0.903	0.248	0.814	0.841	0.792	0.744
GUANGXI	0.882	0.795	0.857	0.864	0.746	0.839	1.000	0.592	0.798	0.570	0.359	0.530	0.699	0.791	0.825	0.895	0.862	0.409	0.338	0.760	0.721	0.529	0.619	0.840	0.414	0.856	0.411	0.763	0.845	0.831	0.850
GUIZHOU	0.582	0.363	0.657	0.773	0.406	0.634	0.592	1.000	0.701	0.115	0.039	0.107	0.253	0.421	0.336	0.520	0.634	-0.006	-0.033	0.561	0.492	0.124	0.269	0.532	0.049	0.509	-0.020	0.479	0.448	0.671	0.397
HAINAN	0.810	0.662	0.903	0.791	0.740	0.929	0.798	0.701	1.000	0.347	0.088	0.346	0.475	0.671	0.651	0.771	0.792	0.115	0.114	0.926	0.778	0.375	0.415	0.831	0.223	0.789	0.111	0.861	0.826	0.741	0.607
HEBEI	0.713	0.718	0.472	0.535	0.774	0.503	0.570	0.115	0.347	1.000	0.822	0.806	0.670	0.829	0.824	0.723	0.729	0.868	0.829	0.439	0.205	0.835	0.903	0.506	0.865	0.616	0.802	0.396	0.568	0.703	0.718
HEILONGJIANG	0.536	0.584	0.273	0.325	0.623	0.271	0.359	0.039	0.088	0.822	1.000	0.869	0.638	0.704	0.584	0.454	0.552	0.894	0.954	0.171	-0.070	0.877	0.897	0.335	0.694	0.445	0.701	0.193	0.257	0.584	0.478
HENAN	0.675	0.803	0.502	0.478	0.777	0.515	0.530	0.107	0.346	0.806	0.869	1.000	0.742	0.816	0.701	0.639	0.657	0.791	0.896	0.421	0.204	0.938	0.933	0.624	0.663	0.663	0.697	0.406	0.481	0.696	0.641
HUBEI	0.803	0.748	0.649	0.689	0.657	0.646	0.699	0.253	0.475	0.670	0.638	0.742	1.000	0.860	0.671	0.792	0.689	0.606	0.657	0.468	0.356	0.685	0.749	0.714	0.557	0.809	0.475	0.505	0.580	0.667	0.812
HUNAN	0.921	0.845	0.771	0.796	0.892	0.795	0.791	0.421	0.671	0.829	0.704	0.816	0.860	1.000	0.857	0.893	0.896	0.696	0.707	0.714	0.458	0.825	0.878	0.773	0.679	0.865	0.619	0.646	0.739	0.857	0.834
INNER MONGOLIA	0.843	0.832	0.727	0.727	0.868	0.728	0.825	0.336	0.651	0.824	0.584	0.701	0.671	0.857	1.000	0.870	0.842	0.695	0.566	0.729	0.639	0.736	0.789	0.737	0.662	0.788	0.669	0.681	0.837	0.814	0.806
JIANGSU	0.942	0.884	0.865	0.905	0.821	0.862	0.895	0.520	0.771	0.723	0.454	0.639	0.792	0.893	0.870	1.000	0.898	0.556	0.475	0.748	0.667	0.629	0.730	0.879	0.638	0.933	0.571	0.715	0.862	0.864	0.950
JIANGXI	0.929	0.787	0.873	0.863	0.843	0.843	0.862	0.634	0.792	0.729	0.552	0.657	0.689	0.896	0.842	0.898	1.000	0.593	0.537	0.763	0.605	0.692	0.769	0.794	0.636	0.866	0.499	0.741	0.804	0.927	0.815
JILIN	0.564	0.603	0.284	0.391	0.621	0.284	0.409	-0.006	0.115	0.868	0.894	0.791	0.606	0.696	0.695	0.556	0.593	1.000	0.858	0.217	0.038	0.824	0.850	0.354	0.823	0.489	0.813	0.219	0.389	0.598	0.586
LIAONING	0.526	0.616	0.292	0.291	0.618	0.315	0.338	-0.033	0.114	0.829	0.954	0.896	0.657	0.707	0.566	0.475	0.537	0.858	1.000	0.201	-0.038	0.866	0.891	0.380	0.754	0.495	0.683	0.235	0.279	0.566	0.506
NINGXIA	0.774	0.694	0.827	0.729	0.820	0.899	0.760	0.561	0.926	0.439	0.171	0.421	0.468	0.714	0.729	0.748	0.763	0.217	0.201	1.000	0.725	0.478	0.482	0.769	0.288	0.747	0.227	0.809	0.848	0.700	0.578
QINGHAI	0.598	0.597	0.808	0.624	0.534	0.728	0.721	0.492	0.778	0.205	-0.070	0.204	0.356	0.458	0.639	0.667	0.605	0.038	-0.038	0.725	1.000	0.199	0.237	0.757	0.161	0.697	0.049	0.789	0.798	0.579	0.549
SHAANXI	0.696	0.771	0.507	0.481	0.823	0.529	0.529	0.124	0.375	0.835	0.877	0.938	0.685	0.825	0.736	0.629	0.692	0.824	0.866	0.478	0.199	1.000	0.927	0.587	0.661	0.621	0.704	0.415	0.495	0.689	0.592
SHANDONG	0.768	0.820	0.561	0.608	0.827	0.585	0.619	0.269	0.415	0.903	0.897	0.933	0.749	0.878	0.789	0.730	0.769	0.850	0.891	0.482	0.237	0.927	1.000	0.633	0.761	0.696	0.752	0.431	0.551	0.808	0.714
SHANGHAI	0.853	0.903	0.916	0.807	0.786	0.917	0.840	0.532	0.831	0.506	0.335	0.624	0.714	0.773	0.737	0.879	0.794	0.354	0.380	0.769	0.757	0.587	0.633	1.000	0.395	0.942	0.347	0.774	0.827	0.797	0.806
SHANXI	0.568	0.584	0.389	0.429	0.585	0.378	0.414	0.049	0.223	0.865	0.694	0.663	0.557	0.679	0.662	0.638	0.636	0.823	0.754	0.288	0.161	0.661	0.761	0.395	1.000	0.549	0.722	0.299	0.467	0.610	0.655
SICHUAN	0.899	0.890	0.907	0.833	0.802	0.903	0.856	0.509	0.789	0.616	0.445	0.663	0.809	0.865	0.788	0.933	0.866	0.489	0.495	0.747	0.697	0.621	0.696	0.942	0.549	1.000	0.418	0.759	0.836	0.835	0.881
TIANJIN	0.507	0.639	0.203	0.408	0.553	0.248	0.411	-0.020	0.111	0.802	0.701	0.697	0.475	0.619	0.669	0.571	0.499	0.813	0.683	0.227	0.049	0.704	0.752	0.347	0.722	0.418	1.000	0.177	0.364	0.553	0.625
TIBET	0.750	0.648	0.847	0.666	0.692	0.814	0.763	0.479	0.861	0.396	0.193	0.406	0.505	0.646	0.681	0.715	0.741	0.219	0.235	0.809	0.789	0.415	0.431	0.774	0.299	0.759	0.177	1.000	0.795	0.710	0.603
XINJIANG	0.817	0.785	0.842	0.781	0.810	0.841	0.845	0.448	0.826	0.568	0.257	0.481	0.580	0.739	0.837	0.862	0.804	0.389	0.279	0.848	0.798	0.495	0.551	0.827	0.467	0.836	0.364	0.795	1.000	0.733	0.774
YUNNAN	0.894	0.812	0.827	0.856	0.813	0.792	0.831	0.671	0.741	0.703	0.584	0.696	0.667	0.857	0.814	0.864	0.927	0.598	0.566	0.700	0.579	0.689	0.808	0.797	0.610	0.835	0.553	0.710	0.733	1.000	0.788
ZHEJIANG	0.866	0.846	0.750	0.840	0.696	0.744	0.850	0.397	0.607	0.718	0.478	0.641	0.812	0.834	0.806	0.950	0.815	0.586	0.506	0.578	0.549	0.592	0.714	0.806	0.655	0.881	0.625	0.603	0.774	0.788	1.000

(a) PC-based job search

	ANHUI	BEIJING	CHONGQING	FUJIAN	GANSU	GUANGDONG	GUANGXI	GUIZHOU	HAINAN	HEBEI	HEILONGJIANG	HENAN	HUBEI	HUNAN	INNER MONGOLIA	JIANGSU	JIANGXI	JILIN	LIAONING	NINGXIA	QINGHAI	SHAANXI	SHANDONG	SHANGHAI	SHANXI	SICHUAN	TIANJIN	TIBET	XINJIANG	YUNNAN	ZHEJIANG
ANHUI	1.000	0.559	0.700	0.654	0.588	0.678	0.670	0.672	0.699	0.621	0.544	0.687	0.736	0.734	0.617	0.784	0.733	0.512	0.674	0.592	0.551	0.600	0.714	0.589	0.598	0.542	0.537	0.675	0.531	0.663	0.522
BEIJING	0.559	1.000	0.817	0.922	0.873	0.914	0.794	0.717	0.720	0.941	0.874	0.923	0.856	0.902	0.915	0.907	0.789	0.908	0.854	0.843	0.762	0.882	0.898	0.944	0.908	0.850	0.917	0.636	0.902	0.831	0.870
CHONGQING	0.700	0.817	1.000	0.844	0.935	0.932	0.937	0.930	0.944	0.931	0.780	0.946	0.930	0.948	0.929	0.928	0.976	0.818	0.838	0.945	0.833	0.863	0.957	0.844	0.905	0.886	0.814	0.835	0.810	0.964	0.639
FUJIAN	0.654	0.922	0.844	1.000	0.860	0.905	0.862	0.796	0.810	0.900	0.726	0.880	0.848	0.917	0.860	0.896	0.848	0.771	0.755	0.829	0.716	0.751	0.874	0.879	0.856	0.777	0.822	0.741	0.954	0.842	0.933
GANSU	0.588	0.873	0.935	0.860	1.000	0.926	0.922	0.876	0.899	0.954	0.818	0.943	0.895	0.914	0.953	0.900	0.918	0.870	0.835	0.964	0.864	0.870	0.946	0.889	0.928	0.911	0.876	0.768	0.861	0.954	0.690
GUANGDONG	0.678	0.914	0.932	0.905	0.926	1.000	0.926	0.886	0.900	0.958	0.837	0.958	0.961	0.964	0.928	0.960	0.925	0.859	0.890	0.921	0.831	0.906	0.961	0.912	0.916	0.887	0.881	0.820	0.864	0.932	0.760
GUANGXI	0.670	0.794	0.937	0.862	0.922	0.926	1.000	0.970	0.967	0.914	0.698	0.900	0.896	0.920	0.886	0.911	0.965	0.734	0.805	0.937	0.774	0.775	0.937	0.821	0.875	0.784	0.801	0.847	0.865	0.931	0.700
GUIZHOU	0.672	0.717	0.930	0.796	0.876	0.886	0.970	1.000	0.980	0.870	0.633	0.855	0.851	0.872	0.839	0.874	0.961	0.671	0.769	0.916	0.761	0.721	0.906	0.751	0.807	0.740	0.733	0.874	0.807	0.924	0.618
HAINAN	0.699	0.720	0.944	0.810	0.899	0.900	0.967	0.980	1.000	0.869	0.646	0.865	0.869	0.877	0.857	0.877	0.970	0.686	0.764	0.927	0.788	0.739	0.903	0.766	0.812	0.787	0.737	0.899	0.790	0.943	0.610
HEBEI	0.621	0.941	0.931	0.900	0.954	0.958	0.914	0.870	0.869	1.000	0.855	0.977	0.930	0.957	0.956	0.958	0.920	0.908	0.887	0.950	0.815	0.916	0.975	0.930	0.952	0.897	0.916	0.734	0.895	0.936	0.777
HEILONGJIANG	0.544	0.874	0.780	0.726	0.818	0.837	0.698	0.633	0.646	0.855	1.000	0.889	0.826	0.822	0.903	0.850	0.731	0.970	0.933	0.817	0.751	0.947	0.869	0.872	0.900	0.821	0.929	0.624	0.693	0.823	0.574
HENAN	0.687	0.923	0.946	0.880	0.943	0.958	0.900	0.855	0.865	0.977	0.889	1.000	0.965	0.967	0.956	0.971	0.923	0.919	0.917	0.939	0.819	0.947	0.981	0.934	0.963	0.919	0.908	0.738	0.845	0.940	0.728
HUBEI	0.736	0.856	0.930	0.848	0.895	0.961	0.896	0.851	0.869	0.930	0.826	0.965	1.000	0.962	0.890	0.964	0.910	0.845	0.892	0.887	0.762	0.925	0.956	0.890	0.911	0.891	0.833	0.757	0.785	0.901	0.687
HUNAN	0.734	0.902	0.948	0.917	0.914	0.964	0.920	0.872	0.877	0.957	0.822	0.967	0.962	1.000	0.921	0.966	0.934	0.848	0.878	0.902	0.788	0.891	0.968	0.893	0.942	0.860	0.862	0.789	0.874	0.917	0.780
INNER MONGOLIA	0.617	0.915	0.929	0.860	0.953	0.928	0.886	0.839	0.857	0.956	0.903	0.956	0.890	0.921	1.000	0.922	0.909	0.928	0.893	0.956	0.866	0.915	0.953	0.911	0.954	0.904	0.907	0.767	0.838	0.938	0.690
JIANGSU	0.784	0.907	0.928	0.896	0.900	0.960	0.911	0.874	0.877	0.958	0.850	0.971	0.964	0.966	0.922	1.000	0.925	0.872	0.925	0.906	0.754	0.910	0.980	0.920	0.923	0.847	0.885	0.779	0.855	0.921	0.766
JIANGXI	0.733	0.789	0.976	0.848	0.918	0.925	0.965	0.961	0.970	0.920	0.731	0.923	0.910	0.934	0.909	0.925	1.000	0.768	0.817	0.940	0.824	0.812	0.947	0.823	0.874	0.835	0.792	0.856	0.818	0.957	0.650
JILIN	0.512	0.908	0.818	0.771	0.870	0.859	0.734	0.671	0.686	0.908	0.970	0.919	0.845	0.848	0.928	0.872	0.768	1.000	0.909	0.864	0.755	0.955	0.893	0.918	0.925	0.872	0.935	0.611	0.754	0.858	0.635
LIAONING	0.674	0.854	0.838	0.755	0.835	0.890	0.805	0.769	0.764	0.887	0.933	0.917	0.892	0.878	0.893	0.925	0.817	0.909	1.000	0.865	0.735	0.935	0.934	0.859	0.906	0.775	0.918	0.712	0.736	0.868	0.604
NINGXIA	0.592	0.843	0.945	0.829	0.964	0.921	0.937	0.916	0.927	0.950	0.817	0.939	0.887	0.902	0.956	0.906	0.940	0.864	0.865	1.000	0.845	0.863	0.955	0.869	0.925	0.873	0.878	0.804	0.837	0.969	0.640
QINGHAI	0.551	0.762	0.833	0.716	0.864	0.831	0.774	0.761	0.788	0.815	0.751	0.819	0.762	0.788	0.866	0.754	0.824	0.755	0.735	0.845	1.000	0.778	0.813	0.751	0.777	0.848	0.739	0.718	0.689	0.839	0.523
SHAANXI	0.600	0.882	0.863	0.751	0.870	0.906	0.775	0.721	0.739	0.916	0.947	0.947	0.925	0.891	0.915	0.910	0.812	0.955	0.935	0.863	0.778	1.000	0.921	0.893	0.918	0.905	0.899	0.646	0.707	0.864	0.588
SHANDONG	0.714	0.898	0.957	0.874	0.946	0.961	0.937	0.906	0.903	0.975	0.869	0.981	0.956	0.968	0.953	0.980	0.947	0.893	0.934	0.955	0.813	0.921	1.000	0.913	0.955	0.868	0.912	0.801	0.860	0.959	0.713
SHANGHAI	0.589	0.944	0.844	0.879	0.889	0.912	0.821	0.751	0.766	0.930	0.872	0.934	0.890	0.893	0.911	0.920	0.823	0.918	0.859	0.869	0.751	0.893	0.913	1.000	0.898	0.871	0.900	0.670	0.853	0.868	0.788
SHANXI	0.598	0.908	0.905	0.856	0.928	0.916	0.875	0.807	0.812	0.952	0.900	0.963	0.911	0.942	0.954	0.923	0.874	0.925	0.906	0.925	0.777	0.918	0.955	0.898	1.000	0.861	0.923	0.706	0.850	0.910	0.706
SICHUAN	0.542	0.850	0.886	0.777	0.911	0.887	0.784	0.740	0.787	0.897	0.821	0.919	0.891	0.860	0.904	0.847	0.835	0.872	0.775	0.873	0.848	0.905	0.868	0.871	0.861	1.000	0.788	0.638	0.715	0.875	0.599
TIANJIN	0.537	0.917	0.814	0.822	0.876	0.881	0.801	0.733	0.737	0.916	0.929	0.908	0.833	0.862	0.907	0.885	0.792	0.935	0.918	0.878	0.739	0.899	0.912	0.900	0.923	0.788	1.000	0.680	0.823	0.874	0.705
TIBET	0.675	0.636	0.835	0.741	0.768	0.820	0.847	0.874	0.899	0.734	0.624	0.738	0.757	0.789	0.767	0.779	0.856	0.611	0.712	0.804	0.718	0.646	0.801	0.670	0.706	0.638	0.680	1.000	0.705	0.852	0.543
XINJIANG	0.531	0.902	0.810	0.954	0.861	0.864	0.865	0.807	0.790	0.895	0.693	0.845	0.785	0.874	0.838	0.855	0.818	0.754	0.736	0.837	0.689	0.707	0.860	0.853	0.850	0.715	0.823	0.705	1.000	0.821	0.917
YUNNAN	0.663	0.831	0.964	0.842	0.954	0.932	0.931	0.924	0.943	0.936	0.823	0.940	0.901	0.917	0.938	0.921	0.957	0.858	0.868	0.969	0.839	0.864	0.959	0.868	0.910	0.875	0.874	0.852	0.821	1.000	0.635
ZHEJIANG	0.522	0.870	0.639	0.933	0.690	0.760	0.700	0.618	0.610	0.777	0.574	0.728	0.687	0.780	0.690	0.766	0.650	0.635	0.604	0.640	0.523	0.588	0.713	0.788	0.706	0.599	0.705	0.543	0.917	0.635	1.000

（B）Mobile-based job search

Notes: The sample period for PC-based job search begins from Jan 2009 to Sep 2018. The sample period for mobile-based job search begins from Jul 2013 to Sep 2018. The lighter orange indicates higher correlation coefficient.

Table 3. Descriptive statistics of job search for each province

Province	Economic Region	Correlation between PC and Mobile-based job search
Tibet	Western	0.738
Chongqing	Western	0.705
Guangxi	Western	0.696
Hubei	Central	0.656
Hainan	Eastern	0.639
Shanghai	Eastern	0.619
Guangdong	Eastern	0.572
Sichuan	Western	0.539
Jiangsu	Eastern	0.533
Qinghai	Western	0.509
Jiangxi	Central	0.451
Beijing	Eastern	0.44
Xinjiang	Western	0.43
Guizhou	Western	0.418
Hunan	Central	0.384
Yunnan	Western	0.373
Henan	Central	0.35
Anhui	Central	0.342
Fujian	Eastern	0.336
Zhejiang	Eastern	0.328
Ningxia	Western	0.305
Liaoning	Northeastern	0.238
Shandong	Eastern	0.218
Gansu	Western	0.183
Inner Mongolia	Northeastern/ Western	0.12
Shaanxi	Western	0.11
Heilongjiang	Northeastern	0.017
Tianjin	Eastern	-0.135
Shanxi	Central	-0.145
Hebei	Eastern	-0.186
Jilin	Northeastern	-0.2
Overall		0.226

Notes: This table reports the correlation between PC-based and mobile-based JSI for each province. Sample period is July from 2013 to Sep 2018. Based on National Bureau of Statistics of China, a part of cities in Inner Mongolia are classified as northeastern region, while the rest is classified as western region.

Table 4. Descriptive statistics of job search for each province

Province	Economic Region	Mean (PC)	Per user	Mean (mobile)	Per 1,000 user
Anhui	Central	54,876.573	2.712	64,878.921	2.475
Henan	Central	101,462.701	3.302	105,856.905	2.670
Hubei	Central	88,651.744	3.687	90,203.429	3.090
Hunan	Central	59,617.650	2.567	58,541.127	2.019
Jiangxi	Central	32,184.897	2.260	36,960.127	1.905
Shanxi	Central	32,306.470	1.943	36,967.825	1.839
Beijing	Eastern	356,591.103	23.900	165,480.317	9.850
Fujian	Eastern	59,171.513	2.496	46,500.222	1.741
Guangdong	Eastern	313,977.333	4.473	358,319.000	4.496
Hainan	Eastern	12,772.573	3.169	17,706.508	3.756
Hebei	Eastern	68,217.923	2.188	78,002.429	1.989
Jiangsu	Eastern	191,755.060	4.723	166,852.381	3.695
Shandong	Eastern	137,023.889	3.322	113,649.841	2.236
Shanghai	Eastern	170,306.350	10.597	132,763.222	7.394
Tianjin	Eastern	89,091.419	10.899	66,377.095	6.772
Zhejiang	Eastern	150,284.607	4.504	94,664.778	2.606
Heilongjiang	Northeastern	49,925.906	3.618	48,515.222	2.740
Jilin	Northeastern	43,641.368	3.948	45,926.063	3.371
Liaoning	Northeastern	98,672.513	4.347	92,855.952	3.391
Inner Mongolia	Northeastern/ Western	18,726.410	1.802	22,190.968	1.717
Gansu	Western	13,092.265	1.534	19,585,540	1.829
Guangxi	Western	21,552.863	1.240	26,742.841	1.238
Guizhou	Western	16,879.906	1.453	22,570,556	1.517
Ningxia	Western	6,731.103	2.591	8,718.444	2.606
Qinghai	Western	5,082.564	1.967	6,539.365	2.054
Shaanxi	Western	70,512.462	4.430	92,242.349	4.740
Sichuan	Western	111,341.632	4.018	143,300.063	4.107
Tibet	Western	3,708.393	3.373	4,181.540	2.896
Xinjiang	Western	10,905.863	0.998	11,999.143	0.937
Yunnan	Western	32,494.872	2.266	43,738.254	2.356
Chongqing	Western	45,362.009	3.565	55,166.587	3.617
Overall		79,578.127	4.126	74,522.826	3.150

Notes: This table reports the average of annual internet searches for job search-related words for each province. All variables are annual observation. Period for PC-based is from 2009 to 2018, while that for mobile-based is from 2011 to 2018. Per capita refers to the average of total annual searches to total population (per 1,000 people). Based on National Bureau of Statistics of China, a part of cities in Inner Mongolia are classified as northeastern region, while the rest is classified as western region.

Table 5. Weekday, weekend, and holiday effects

	(1)	(2)	(3)	(4)	(5)	(6)
	PC-based			Mobile-based		
Monday	0.694***	0.638***		0.147***	0.146***	
	(0.015)	(0.014)		(0.006)	(0.007)	
Tuesday	0.692***	0.636***		0.146***	0.145***	
	(0.015)	(0.014)		(0.006)	(0.007)	
Wednesday	0.676***	0.620***		0.139***	0.137***	
	(0.015)	(0.014)		(0.006)	(0.007)	
Thursday	0.686***	0.630***		0.131***	0.130***	
	(0.015)	(0.014)		(0.006)	(0.007)	
Friday	0.597***	0.542***		0.105***	0.104***	
	(0.015)	(0.015)		(0.006)	(0.007)	
Saturday	0.056***			0.003		
	(0.016)			(0.007)		
Sunday		-0.056***			-0.003	
		(0.016)			(0.007)	
Weekend			-0.641***			-0.134***
			(0.009)			(0.004)
Holidays	-1.182***	-1.182***	-1.181***	-0.306***	-0.306***	-0.306***
	(0.024)	(0.024)	(0.024)	(0.012)	(0.012)	(0.012)
Constant	3.370***	3.426***	4.039***	5.768***	5.770***	5.902***
	(0.014)	(0.013)	(0.008)	(0.005)	(0.006)	(0.0044)
R-squared	0.508	0.508	0.507	0.696	0.696	0.696
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	110,360	110,360	110,360	59,458	59,458	59,458

Notes: The dependent variable in all regressions is the log of job search index (JSI). Independent variables include weekday, weekend, and holiday dummies. We control for province and week fixed effects. Standard errors are clustered at the province level and corrected for heteroskedasticity. The clustered standard errors are in parentheses. The sample period for PC-based JSI is daily observations from 1st July 2009 to 30th Sep 2018, while that for mobile-based JSI is from 1st Jan 2009 to 30th Sep 2018. *, **, *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6. Job search over business cycle

	(1)	(2)	(3)	(4)	(5)
	Overall	Central	Eastern	Northeastern	Western
Panel A: PC-based					
Provincial GDP	0.655***	0.686*	2.298***	-0.683***	1.205***
	(0.146)	(0.374)	(0.311)	(0.151)	(0.382)
Constant	0.019**	0.005	0.122***	-0.193***	-0.002
	(0.008)	(0.012)	(0.010)	(0.020)	(0.020)
Time FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.688	0.828	0.773	0.882	0.532
Observations	837	162	270	108	297
Panel B: Mobile-based					
Provincial GDP	-0.146	0.602**	-0.249	-0.050	-1.012**
	(0.189)	(0.242)	(0.482)	(0.173)	(0.431)
Constant	-0.006	0.050***	0.063***	-0.156***	-0.072***
	(0.012)	(0.013)	(0.015)	(0.035)	(0.024)
Time FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.779	0.785	0.830	0.918	0.728
Observations	279	54	90	36	99

Notes: The dependent variable in all regressions is the job search index (JSI). The HP (Hamilton) detrended represents JSI constructed based on HP (Hamilton) filter. Main independent variable is provincial GDP. We control for province and time fixed effects. Standard errors are clustered by both province and time and are reported in parentheses. The sample period for PC-based is from Jan 2009 to Sep 2018. The sample period for mobile-based JSI begins from Jul 2013 to Sep 2018. All are quarterly observations. *, **, *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.