**Labour Market Insights: the Power of Internet-Based Data**

Nuarpear Warn Lekfuangfu[1], Voraprapa Nakavachara and Paphatsorn Sawaengsuksant

Preliminary: 1 September 2016

(Please do not quote)

## Abstract

In this project, we will conduct a series of research exercise to demonstrate how selected web-based data sources can provide additional insights for labour market analysis, beyond what conventional government-conducted surveys can offer. To begin, we will briefly explore the role of Google Trends for forecasting and nowcasting labour market indicators in Thailand's context. Then, we exploit web-based data from selected job-boards and resume postings under Thai domain to provide some insights on job vacancy statistics, labour market mismatch between required skill *vis-a-vis* attained skill at occupation level and the gap between reservation wage and productivity. We also use this dataset to investigate labour market discriminations using separate perspective of firms and job seekers.

**Keywords:** Internet job search, Employment outcomes, Nowcasting, Google Trend, Mismatch

**JEL Codes:** J2, J64

[1] Lekfuangfu and Nakavachara are lecturers at Chulalongkorn University, Faculty of Economics. Lekfuangfu is a research associate at the Centre for Economic Performance (LSE). Sawaengsuksant is at the Bank of Thailand. The authors would like to thank Donn Tawanpitak and Pakawin Mansukmanee for their excellent assistance with the data. The results and findings reflect the personal view of the authors and do not reflect the views of their professional institutions.

## I. Background of conventional labour market indicators of Thailand

In Thailand, there are a number of sources that help provide meaningful insights to labour market conditions of the economy. To track labour market behaviours at the macro-level, policy makers conventionally rely on the statistics drawn from the monthly Labour Force Survey (LFS) to depict the states of employment equilibrium of the country- more famously the unemployment rate. To gain fuller comprehension of the demand and the supply side of the market separately, more statistics are drawn from various sources. For the demand side, the Business Sentiment Index (BSI) and the Labour Demand of Establishment Surveys (LD) are often referred to. In both datasets, only selected sectors are surveyed[2]. In addition, the Ministry of Labour (MoL) also reports the aggregate number of vacancies reported by firms that directly seeking via the channels provided by the Department of Employment. For the status of labour supply in Thailand, the MoL also reports the monthly total dismissals nation-wide. However, most of the MoL data will over-represent workers and employers within the Social Security System[3].

In detail, the National Statistical Office of Thailand (NSO) provides the indicator of the unemployment rate on a monthly basis. Basically the unemployment rate is the number of unemployed workers (not working but looking for work or available for work) divided by those in the labour force. The NSO starts collecting the survey information at the beginning of each month then processes, checks, analyses and announces it on the 5th day of the following month.

Figure 1.1 shows Thailand's unemployment rate from 2004 to early 2016. It appeared that the unemployment rate could only slightly pick up the shock from the great financial crisis that hit Thailand in 2009. However, it was not able to precisely pick up the shock from the 2011 flood.

It has been discussed in the literature that shocks affecting the formal sector of the labour market may not affect the unemployment rate. This is because the informal sector of the

[2] The Business Sentiment Index (BSI) survey privates companies and calculate diffusion index of current hiring situation and hiring plans in next 3 months. The Labour Demand of Establishment Surveys (LD) survey medium-to-large enterprises in 8 sub-sections of manufacturing and some services. For the latest survey in 2013, it covered establishments of 6 employees or more for selected sectors. This counts as 31,856 establishments nation-wide. The sectors covered in the survey are: commercial, service, construction, manufacturing, land transportation, tourism and hospitals.

[3] Currently, the MoL also produces its own labour market indicators, under the name of Labour Market Warning. This index includes the statistics mentioned above and also other industries' macro-level indicators. Very interestingly, it includes the total number of job advertisement. But this is also the number taken from print outlets. Given the downward trend of prints, this number is believed to hugely under-represent the actual total job ads at a point in time.

labour market may have absorbed the workers dismissed from the formal sector employment. In addition, the employers may respond to shocks by reducing the employees' work hours rather than dismissing them. However, the unemployment rate is still worth paying attention to since there could still be shocks that affect the informal sector or there could still be situations where the informal sector could not or could only partially absorb the shocks in the formal sector.

To clearly observe the shocks that affect the formal sector of the labour market, one could look at Thailand's registration for unemployment. The Social Security Act B.E. 2533 (1990) mandated that workers employed in the formal sector (excluding public officials) be insured under this Act (per Section 33). One of the benefits that the workers get is that if they ever become unemployed (either because they have been dismissed or because they quit), they are eligible to receive unemployment benefits (at the rate of 50% of their previous wage for not more than 180 days in the case of dismissal; and at the rate of 30% of their previous wage for not more than 90 days in the case of resignation). However, in order to receive such unemployment benefits, workers are required to register their unemployment at the Department of Employment, Ministry of Labour within 30 days of becoming unemployed.

Figure 1.2 shows Thailand's registration for unemployment from 2004 to early 2016[4]. Two types of the unemployment registration graphs are shown. The first one is for the case of dismissal and the second one is for the case of resignation. The Department of Employment (Ministry of Labour) is responsible for these monthly data series. There are no clear guidelines on how many days (after the end of the month) that these data would be released to public. However, the authors have observed that these data are usually released on the Ministry of Labour's website by the end of the following month (i.e., approximately 30 days lag)[5].

It appears that both graphs were able to pick up the 2009 financial crisis at a significant amount. In addition, for the dismissed workers case, the graph was also able to pick up the shock from the 2011 flood. Note that these two graphs involve only formal sector employees under the Social Security Act (and thus are eligible to claim for the unemployment benefits when become unemployed). As of June 2016, there are 10,377,038 workers insured under Section 33 and 2,205,170 workers insured under section 40 of the Act. According to the NSO's

[4] The data that are available to public started from July 2004.
[5] The authors referred to this website http://warning.mol.go.th/index.php/page_detail?id=20.

Informal Employment Survey for 2015, there were about 16.9 million workers employed in the formal sector and about 21.4 million workers employed in the informal sector.

In this research, we acknowledge that the conventional sources of information, in particular the LFS, remains significantly valuable to provide insights to labour market conditions of Thailand. One notably advantage of this survey is that it has been conducted for a long period of time- therefore this consistent continuation allows researchers and policy makers to understand the dynamic of Thai labour markets many decades back. Moreover, as a nationally representative survey, one can constructively draw lesson learnt from statistical analysis using this dataset to the national platform.

However, it is worth recognising that a collection of statistics from the LFS, especially the employment status as well as salaries of workers, is the outcomes of the demand and supply interactions. That is, these are average salary or unemployment rate at the equilibrium. While it, in general, tells a decent story of labour market, it cannot un-wrap the mechanisms within the markets. A number of advanced macro-economic models are generally embedded with a feature of labour market matching model (Mortensen and Pissarides, 1994; Shimer, 2005). This model requires information from each separate side of the same markets. While the LFS can provide the unemployment rate (u), it lacks information at all on the demand side – the vacancy rate. We note that one may use the LD or the MoL's vacancy report as the proxy. But, one must also note that these numbers are not only generated at completely different times, they are taken from different economic sectors and most of all, do not represent the same matched sampling pools.

In many other countries, there exists a number of ways in which the statistics for the crucial "labour market tightness (v/u)" can be calculated more properly[6]. For example, in France, researches make use of a matched employer-employee dataset- whereby the samplings are taken from the administrative data[7]. In the US, the Bureau of Labor Statistics produces, on top of the LFS equivalent (Current Population Surveys), the Job Openings Labor Turnover

[6] Labour market tightness composes of two key features of the Beveridge Curve- vacancy and unemployed. When an economy is expanding, v generally gets bigger and u shrinks. Therefore the tightness (calculated as v/u) will get smaller, and vice versa. Historically, the US data shows a negative relationship between v and u. However, this relationship breaks down in the decent decade- where v may get bigger but not much change of u.

[7] For example, the D´eclaration Annuelle des Donn´ees Sociales (DADS) (DADS) is an administrative database of matched employer-employee information collected by the INSEE (Institut Nationale de la Statistique et des Etudes Economique). The data are based on the mandatory reports, filed by employers, of the gross earnings of each employee in compliance with French payroll taxes. All wage-paying individuals and legal entities established in France are required to file payroll declarations; only individuals employing civil servants are excluded from filing such declarations. But this dataset lacks vacancies.

(JOLTS) as its complementary monthly survey for the vacancies from a sample of nonfarm establishments of entities in the 50 states[8]. JOLTS data are collected for total employment, job openings, hires, quits, layoffs and discharges, other separations, and total separations. However, it is worth notice that, as a survey, JOLTS sample under-represents sectors or enterprises with high exit rate (fail within the first year) (BLS, 2016). This is mainly due to the time-consuming process of any surveys[9].

More recently, the US's Conference Board Help Wanted *Online* Data Series (HWOL) has been in the focus (Barnichon, 2010, 2015; Sahin, 2011). This is an updated series of the Conference Board Help Wanted Series, which report total number of job ads from print media. The attention directed to HWOL is a response to the recognition of the growth of digital activities in the US in the recent decades. Unlike the old version, HWOL consists of job ads from both the prints as well as from the online job boards across all US states. And unlike JOLTS, HWOL reports the actual total number of job vacancies of the entire pool of advertised job openings across number of industries, in a timelier manner[10]. Figure 1.3 tracks the time trends of US unemployment rate (CPS) and the job vacancies from JOLTS and HWOL over time.

Our research aims are to explain and understand behaviours of labour markets in Thailand from alternative datasets. These chosen datasets in this research share three features. First, all datasets reflect the state of labour markets in Thailand in real-time. Second, they are virtually costless to collect and they are user-generated information. And third and foremost,

[8] The JOLTS survey design is a stratified random sample of 16,000 nonfarm business and government establishments. The sample is stratified by ownership, region, industry sector, and establishment size class. The establishments are drawn from a universe of over 9.1 million establishments compiled by the Quarterly Census of Employment and Wages (QCEW) program which includes all employers subject to state unemployment insurance laws and federal agencies subject to the Unemployment Compensation for Federal Employees program

[9] Moreover, Jobs to be filled only by internal transfers, promotions, demotions, or recall from layoffs are excluded. Also excluded are jobs with start dates more than 30 days in the future. The separations level is the total number of employment terminations occurring at any time during the reference month, and is reported by type of separation—quits, layoffs and discharges, and other separations.

[10] The HWOL program contains data collected from over 16,000 online job sources including traditional job boards, corporate boards, and social media sites. Internet job sources that are aggregators (i.e. only scrape ads from other sources and provide no unique ads) are identified and removed from active collection in order to eliminate a major source of duplication in counting online ads. New job sources are identified using independent research and recommendations from industry sources across the U.S. This process results in periodic updates to the HWOL coverage. Job sources that cover smaller niche markets are also included in HWOL; however, smaller local job sources in an area with a limited number of ads may not be targeted for collection. The coverage are occupations to 2 digits, at regional and Metropolitan (MSA) level. (*https://www.conference-board.org/data/helpwantedonline.cfm)*

these datasets reflects the growing shift towards digital activities and economies of markets in Thailand and elsewhere (Autor, 2001; Nakamura et al. 2009, Savankulov 2010).

As the world is becoming more digitized, more and more people rely on internet for their everyday activities. Currently there are 38 million internet users in Thailand which account for 56% of the total population (as of January 2016). The internet penetration rate for Thailand, defined as percentage of the population using internet divided by the total population, has increased significantly over the years, from 26% in 2014, 37% in 2015, to 56% in 2016. The penetration rate is currently 29% for Asia Pacific and 48% worldwide[11].

Based on the Information and Communication Technology Survey (ICT thereof) conducted by the National Statistical Office of Thailand (NSO) in 2015[12], there were 1,062,208 people (aged 15 and older) reported using internet for job searches and application submissions. Most of them were among 20-29 years old (about 42%). Although the number of online job searchers may seem small compared to the total population, this number is not small compared to the number of unemployed people in quarter 1 of 2016 which was 370,000 (see Table 1.1). In addition, the fact that there were more people searching for jobs online compared to the number of people unemployed suggested that there could be quite a number of on-the-job searches.

Globally, the internet penetration rate has increased. Concurrently, there is also an increasing number of research papers that used internet-based data to explain various dimensions of the economy, and of our particular interest, to analyse the labour market. Lenaerts, Beblavy and Fabo (2016) review the literature on how various types of internet-based data such as Google Trends, LinkedIn, Facebook, Twitter, and Glassdoor Survey have been used or discussed in labour-related research. Antenucci, *et al* (2014) created the University of Michigan Social Media Job Loss Index based on job-related messages that people posted on their Twitter accounts. This index appeared to outperform the existing forecast for US initial claims for unemployment insurance. Marinescu (2016) utilizes the data from CareerBuilder.com, a large US internet job board, to analyse the impact of the increased unemployment benefits on job applications and job vacancies. For the Chinese labour market,

[11] Sources: Global Digital Statistics 2014, We Are Social SG; Digital Social & Mobile in 2015, We Are Social SG; Digital in 2016, We Are Social SG.

[12] Note that the internet penetration rate reported by this survey was 39.3% for 2015. In the previous paragraph we utilized the statistics by *We Are Social* since this source collected such information over time whereas the ICT survey only has the information for 2015. We acknowledged that the inconsistency between the two sources may be due to how they collected and processed their data.

Khun and Shen (2013) utilized the data from Zhaopin.com, the third largest internet job board in China, to analyse gender discrimination.

In additional, a collection of study look at the benefits of internet as a job search tool. For applicants, they find that job-seekers are more likely to get employed or have shorter unemployment duration if they also searched on internet[13] (see for example Freeman, 2000; Bagues and Labini, 2007, Stevensen, 2008 and Suvankulov 2010). The study by Kuhn and Skuterud (2004) are the exception. They use the US CPS Computer Supplements and find a relationship between shorter duration of unemployment and using internet as a tool. However, because of the effect disappeared after controlling for many observed individual characteristics, they eventually concluded that internet makes no difference. Fountain (2005) look at the same CPS but analysed the time periods separately. She finds that an early adopter of internet as a job search tool did benefit from using this mode but not the same is found for later cohorts.

Not only has the Internet revolution changed the behaviours of job seekers, it has also altered significantly the recruitment methods of many firms. A survey of the Fortune Global 500 shows that the year 2000, all of these top companies made use of e-recruiting services (Nakamura et al 2009). In comparison to print ads, online platforms have reduced the financial costs of filling vacancies and increased the reach to the public most significantly (see Table 1.1 and 1.2 in Nakamura et al 2009).

We will begin by discussing our use of information generated via the biggest search engine, Google, to deepen our understanding of Thai labour market at the macro-level. Then, we will discuss another user-generated database from online-platforms of labour markets. With this dataset, we will be able to analyse labour market behaviours from the supply side (applicants) and the demand side (job ads) of the markets.

## II. Nowcasting with Google Trend

Some of the researchers were particularly interested in using the internet-based data to nowcast the existing economic indicators. Choi and Varian (2009) explain nowcasting as "predicting the present." While forecasting means using the previous period's data to predict

[13] In detail, Suvankulov(2010) uses German EOSP, Korea's KLIPS and US LFS and find a positive associate of 7% and 12.7% respectively more likely to being employed the next year.

the following period's economic indicators, nowcasting simply means using the current period's data to predict the current period's economic indicators.

The fundamental of nowcasting is that some data (either the traditional type or the internet-based type) come out more frequently or come out sooner than the actual official economic indicators. In the case of traditional data, researchers have tried to nowcast GDP, which is a quarterly indicator the comes out way later than the end of the quarter, using other data series relating to manufacturing, trade, employment, etc. as inputs to their nowcast models. Currently both the Federal Reserves of Atlanta and the Federal Reserves of New York have published their US GDP nowcast updated frequently on a weekly basis or as soon as the new data become available (Higgins, 2014). Other countries' GDPs have also been nowcasted in the literature (Bell, Co, Stone, and Wallis, 2014; Hara and Yamane, 2013; Bragoli and Modugno, 2016; Luciani, Pundit, Ramayandi and Giovanni, 2015; and Modugno, Soybilgen and Yazgan, 2016). Other researchers also have tried to nowcast inflation under this concept (Knotek and Zaman, 2014; and Giannone, Reichlin and Small, 2005).

The other branch of the nowcast literature utilized the internet-based data which are available real-time. Specifically, the internet-based data used in many of these studies are Google Trends, Google's search volume index indicating how often a term or a phrase has been searched by internet users relative to other terms or phrases over a period of time. The economic indicators that have been nowcasted in the literature include retail sales, car sales, home sales, travel volume, price, and of our particular interest, unemployment rate and initial claims for unemployment insurance (Choi and Varian, 2009a, 2009b; Askitas and Zimmermann, 2009; Suhoy, 2009; Carriere-Swallow and Labbe, 2013; Fonduer and Karame, 2013; Vincente, Lopez-Menendez, and Perez, 2015; and Seabold and Coppola, 2015).

In Thailand, to the best of our knowledge, there is currently no research paper using internet-based data to analyse the Thai labour market. Therefore, our paper aims to perform such analysis. Specifically, this section of the paper will use Google Trends to nowcast two labour market indicators namely (i) unemployment rate and (ii) registration of unemployment (dismissed workers). It is also our intention to promote the use of internet-based data in economics research and urge the policy makers to pay attention to such data.

### 2.1. Google trends

The internet-based data used in our analysis are Google Trends. Google Trends report a search volume index of a keyword. The index tells how often the keyword was searched relative to the total number of searches from the same time/location, normalized to be 0 to 100 over the selected time period. Such search volume index is available from January 2004. The period of the data used for our unemployment analysis is from January 2004 to June 2016. The period of the data used for our registration for unemployment (dismissed workers) analysis is from July 2004 (earliest data point available for public) to June 2016.

Since Google Trends do not provide the actual search volume, we utilized Google AdWords' Keyword Planner to find out the average monthly search of a keyword. If the average monthly search of a keyword is low, then such keyword may not be a good representation of the actual search traffic by internet users. Therefore, we also checked the average monthly search when selecting our keywords[14].

The selected keywords used for nowcasting unemployment rate are "หางาน" (Thai word for "job search") or "สมัครงาน" (Thai word for "job application"). It is reasonable to believe that people may search those keywords when they feel that they need to look for a new job. And for nowcasting registration for unemployment (dismissed workers), the selected keyword is "ประกันสังคม" (Thai word for "social security"). It is reasonable to believe that people may search this keyword if they have recently been dismissed and want to find out about how they can receive their unemployment benefits that they are entitled to under the Social Security Act.

### 2.2. Nowcasting unemployment rate and results

Our keyword "หางาน" (job search) has an average monthly search of 368,000 (as of June 2016, the number is rounded by Google AdWords' algorithm) and most of the searches in June 2016 came from Bangkok (60.3%). Our keyword "สมัครงาน" (job application) has an average

[14] Note that it is unclear to us which algorithm was used by Google AdWords' Keyword Planner in calculating the average monthly searches of a keyword. There appeared to be only some selected rounded numbers displayed. However, this should be an acceptable indicator of how often the keyword has actually been search on a given month.

monthly search of 301,000 (as of June 2016, the number is rounded by Google AdWords' algorithm) and most of the searches in June 2016 came from Bangkok (59.1%).

Plotting the Google Trends of the selected keywords against the unemployment rate it is shown that all graphs appeared to move in similar directions (see Figure 2.3)[15]. As discussed earlier, the monthly unemployment rate is available on the 5th day of the following month (lag 5 days). However, Google Trends are available real-time on a weekly frequency[16]. Table 2.1 summarises the variables used in this analysis.

The base model is the AR(1) process as follow:

$$y_t = a + b_1 y_{t-1} + \varepsilon_{it} \qquad \text{(Equation 1)}$$

where $y_t$ is the unemployment rate of the current month. $y_{t-1}$ is the lag term. $\varepsilon_{it}$ is the error term. The Dicky Fuller test was conducted on the unemployment rate data and we could reject the Null of a unit root (non-stationary) process at 1%. Thus the AR(1) model can be used.

The models with Google Trends are as follow:

$$y_t = a + b_1 y_{t-1} + c_1 GSearch_t + \varepsilon_{it} \qquad \text{(Equation 2)}$$

$$y_t = a + b_1 y_{t-1} + c_2 GApply_t + \varepsilon_{it} \qquad \text{(Equation 3)}$$

$$y_t = a + b_1 y_{t-1} + c_1 GSearch_t + c_2 GApply_t + \varepsilon_{it} \qquad \text{(Equation 4)}$$

where $GSearch_t$ is the Google Trends for "หางาน" (Job Search). $GApply_t$ is the Google Trends for "สมัครงาน" (Job Application).

To compare the forecast accuracy among the models, we use different types of prediction errors namely, Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The regression results are shown in Table 2.2.

It appeared that including Google Trends in the regression improved the overall fit of the model (higher Adjusted R2 for Equations 2, 3, and 4 compared to Equation 1). Google

[15] Both of the Google Trends were downloaded separately on 11 July 2016 on a weekly frequency. The Google Trends data of the first week of the month (the week containing the first Sunday of the month) were used in our unemployment nowcasting analysis (as plotted in the Figure 2.3 above).

[16] At the time of our analysis, Google Trends provide weekly data as default. However, later on, Google Trends provide weekly data as default if the requested period is 5 years or less and provide monthly data for longer period. Nevertheless, there appears to be an algorithm in which weekly data can be obtained.

Trends' coefficients are significant in all models. Comparing the 3 types of the prediction errors, we can conclude that Equation 4, which includes both of the Google Trends ("หางาน" (Job Search) and "สมัครงาน" (Job Application)), outperforms the rest of the models.

### 2.3. Nowcasting Registration for Unemployment (Dismissed Workers)

Our keyword "ประกันสังคม" (job application) has an average monthly search of 301,000 (as of June 2016, the number is rounded by Google AdWords' algorithm) and most of the searches in June 2016 came from Bangkok (52.1%) (see Figure 2.4). Figure 2.5 plots the Google Trends of the selected keyword against the registration for unemployment (dismissed workers). The figure shows that the graphs appeared to move in similar directions for most of the time period.[17]

As discussed earlier, the monthly registration for unemployment (dismissed workers) is available by the end of the following month (lag 30 days). However, Google Trends are available real-time on a weekly frequency.[18] Table 2.3 summarises the variables used in this analysis.

The base model is the AR(1) process as follow:

$$y_t = a + b_1 y_{t-1} + \varepsilon_{it} \qquad \text{(Equation 5)}$$

where $y_t$ is the number of the people registered for unemployment (dismissed workers) of the current month. $y_{t-1}$ is the lag term. $\varepsilon_{it}$ is the error term. The Dicky Fuller test was conducted on the registration for unemployment (dismissed workers) data and we could reject the Null of a unit root (non-stationary) process at 5%. Thus the AR(1) model can be used. The models with Google Trends are as follow. Let $Gss_t$ be the Google Trends for "ประกันสังคม" (Social Security).

$$y_t = a + b_1 y_{t-1} + c_1 Gss_t + \varepsilon_{it} \qquad \text{(Equation 6)}$$

[17] The Google Trends were downloaded on 11 July 2016 on a weekly frequency. The average of the first 3 weeks of each month's Google Trends data were utilized in our registration for unemployment (dismissed workers) nowcasting analysis (as plotted in the figure 2.5).
[18] At the time of our analysis, Google Trends provide weekly data as default. However, later on, Google Trends provide weekly data as default if the requested period is 5 years or less and provide monthly data for longer period. Nevertheless, there appears to be an algorithm in which weekly data can be obtained.

Similar to the unemployment analysis, to compare the forecast accuracy among the models, we use different types of prediction errors namely, Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The regression results are shown in Table 2.4.

It appears that including Google Trends in the regression improved the overall fit of the model (higher Adjusted R2 for Equation 6 compared to Equation 5). Google Trends' coefficient is significant in Equation 6. Regarding the prediction errors, only under MSE, the model with Google Trends (Equation 6) outperformed the base model (Equation 5).

For the registration for unemployment (dismissed workers) nowcast, we note that the model with Google Trends outperforms the basic model only under some criteria but not all. This could be due to the fact that the people searching the keyword "ประกันสังคม" (Social Security) may also refer to other types of Social Security benefits that are not unemployment benefits. (These other types of benefits include compensation for sickness, injury, childbearing, etc.) We have tried to use the keyword "ประกันสังคมว่างงาน" (Social Security for Unemployment Benefits) but the average monthly search volume was only 8,100 compared to the average monthly search volume for "ประกันสังคม" (Social Security) of 301,000.[19] Therefore, the monthly search volume may not be large enough to be a meaningful input to the nowcast model. In addition, it appears that other types of social security benefits do not seem to correlate with the economic shocks the same way the unemployment benefits would (i.e., during economic shocks there could be many people being dismissed from their jobs and they would want to receive their social security unemployment benefits). Therefore, we decided that the keyword "ประกันสังคม" (Social Security) would still be the most appropriate keyword although with some noises and limitation.

**III. Understanding two sides of the same coin with data from internet-based job portals**

Turning from using internet-based, user generated data to understand macro-level labour market economics in real-time, we now focus on another type of data that is also made available as a result of the growing role of the Internet in our society. As mentioned in the

[19] Putting it in the regression also does not result in significant coefficient.

previous section, the existing survey datasets of labour markets in Thailand, in particular the LFS, is very valuable but yet remain limited for a number of reasons. First, they take time to prepare, collect and report. Therefore, there is an element of time lag in the dataset. Second, the respondents are propped to reply to the set of questionnaire, and most of the time in person. This leaves a lot of room for under-reporting non-favourable behaviours and over-reporting the favourable ones. In the case of employment status, individuals may be more inclined to over-report positive employment status and under-report unemployment duration or irregular market behaviours. Third, there is no equivalence of data (same markets, with same time frequency) from the demand side to match and compare.

We follow the recent literature within the field, and utilise the power of computer programming in order to obtain an alternative dataset for understanding labour market conditions in Thailand. We turn to the data extracted from Thai's internet-based job portals and obtain two main types of datasets- the job ads and the resume posts. These datasets contain a number of desirable element. First, the datasets are instant reflection of the contemporary market at the time. (as it is generated instantly by the users and can also be instantly extracted). Second, the data is obtained impersonally (not face-to-face) for self-report information in a real stake situation[20]. Third, for second users, these datasets are costless. And most importantly, the information from the demand side (vacancies posted by firms) and the supply side (resume posted by potential job seekers) are at the same time of access and broadly representing the same market. Therefore, this will allow us to take a close look at labour market tightness as well as other interactions within the labour markets which come from two sides of the same coin.

In this paper, we will make use of two levels of data taken from internet-based job portals in Thailand. First, we will use to aggregate-level data for Thailand's biggest job portals to trace the time-varying trends of different characteristics of Thai labour markets. Second, we will turn to a micro, individual-level data to complement our analysis for Thai labour market at a point in time.

[20] Unlike survey questions, firms which post job ads on these job portals are committed financially to the cost charged by these job portals. For applicants, their probability of actual hiring depends strongly on the information they give to these websites. Therefore, they will be less likely to mis-report their personal information than in an equivalent survey question.

### 3.1 Description of the dataset and variables contained (aggregate data)

The aggregate data we use is the publicly available data that the job portal website (Job Portal No. 1) has summarised on its pages. By locking on to the main pages, we can extract total number of job ads as well as of resumes presently on the site at the time of access. For the job ads, these numbers are also disaggregated by occupational categories (41), industry sectors (41) and locations (province-level). Similarly, numbers of resume are also grouped into occupations, fields of study (54), levels of qualification (3), age groups (5) and locations seeking work (province-level).

To obtain past information that no longer available on the present page, we use the service provided for free by an internet-archive web (http://archive.org/web). This internet-archive website allows us to "crawl" back to previous dates on a website of interest and collect any information that the internet-archive may have captured. For the Job Portal No. 1, we are able to crawl back and consistently collect relevant aggregate information from 2001 (the job portal website was firstly launched in 2000). Even if the dates for each year we crawled are not exactly matched, we are able to access and collect the data for at least once for each half of a calendar year. In sum, we are able to obtain the aggregate data on jobs and resume from the Job Portal No. 1 for 31 periods until August 2016.

### 3.2. Description of the dataset and variables contained (micro data)

The micro-data we use here are the universe of unique job advertisements and of unique resumes of job searchers posted on a top online internet-based job portal (renamed as the Job Portal No. 2) during June 2016 to July 2016[21]. To collect the data, we made use of an automatic web crawling programme that searched separately for jobs and resumes posted on this website on a given period. As a result, we obtained all information of each job and each resume that are made publicly viewable. In details, there is a stark different between the nature of job advertising and resume posting. For firms, posting a job ads is costly. In general, a firm pay the website company for a bundle of service, including the duration of ads or the volume of ads and the total number of resumes the firm can view. (On average, the charge is of a minimum of one month of job-posting period.) In contrast, individual job seekers can put up their resume

[21] When conducting a google search with the key word "looking for a job" (in Thai), this website is the top ranking site.

profile free of charge. Most job portals based in Thailand require the posting firm to also submit its official business registration in order to get verification. And all of the job portal website we have visited altogether bar job recruiting for direct sale, peer-to-peer marketing and insurance agents.

We collect the dataset at the individual job and individual resume level. Recall that our data is gathered from a public access data- therefore it does not contain any personal identification (especially the national identification number) or name of applicants. For the firm side, even though we do not have the firm's business registration number, we observe the name and address of these firms.

Note that in a given period our programme crawled the website, there exists jobs and resumes which had been posted before the dates of data access. For job ads, we also collect the date the job was last updated. Recall that a job post will normally remain in the database at least one month- even if the job may have been filled earlier. Therefore, our database of the job-post at the time of access will contain (a) jobs that still exist and (b) jobs that no longer exist but still within the period the firm has paid for. In total, we have 41,312 unique jobs ads during June-July 2016.

Since posting for a resume is free, our task to distinguish the active job-seekers from inactive ones is more challenging. To deal with this, we make use of the date at which a particular resume is last updated. In additionally, we acknowledge that there is a collection of resumes which are created but were not set as publicly searchable. Our database does not contain this type of resume. In total, we gather, within our chosen time period, over 80,000 resume for June-July 2016.

Further investigation (by reading contents on web forum) indicate that our chosen job portal websites represent a non-farming, formal, middle-to-low level sector of the labour market in the country. (Our further investigation, using web forum, indicates that there is a clear market segmentation among Thailand-based job portal websites- between the sites that match top-level jobs vs those that match middle-low level jobs.)

Occupation and industry categories are well indicated in both job ads and resume. The website has 73 tailored occupation categories as a drop-down option (with the addition of "others" and "intern/part-time" categories). Once matched with the International Standard of Occupation Code (2 digit-ISCO version 4, 2008), we aggregate these detailed occupations into

28 groups. As expected, the job ads do not contain any vacancies for agricultural-related (coded 9 of ISCO) and very few elementary service occupations (coded 8 of ISCO).

From the job ads, useful variables include number of vacancies per post (average 3 vacancies each), salary offered (47% of data), minimum work experience requirement (56%), education qualification (90%) as well as total number of views a job ads received between the date posted and the date we accessed the data.

Note that by law, firms based in Thailand are eligible to specify (a) gender, (b) age and/or (c) marital status of candidates they look for. (In fact, firms can also ask for military training background, fertility status or even a photograph of job applicants.). Our dataset also contains some of these insightful characteristic of Thai labour markets. Of all job ads, 75% have target age range, 37% have gender-targeted (of which 14.7% target female applicants and 18% target males).

### 3.3. Data representativeness

As pointed out in Khun and Shen (2013), one can think of many reasons to suspect that our job portal sample is unrepresentative of the nation-wide actual vacancies and pool of job seekers. Despite the growth of the Internet and growing trend of digital-based economy, not all vacancies are captured in the internet-based job portals. We acknowledge that our database may not capture jobs which are (a) internally promoted, (b) relatively short vacancy spell, (c) upper-level and (d) of agricultural sector. In addition, even if we observe some jobs of the public sector, our sample under-represents vacancies from this sector[22].

Given that the pool of Internet users are younger, more educated and of course, more technologically apt than the population, our resume sample reflects the same story (Kuhn and Shen, 2013; Kurekova, 2014; Mang, 2012)[23]. To understand the degree of representativeness of our internet-based dataset (both the aggregate version and the micro-data), we map a number of characteristics from our sample with two surveys of the National Statistical Office- which are nationally representative.

[22] In Kuhn and Shen (2013), they find that the pool of job ads from their Chinese job portal over-represented entre-level jobs, jobs in expanding, jobs with longer vacancy duration and of industries with high turn-over.
[23] As discussed earlier, Thailand's Internet penetration rate is 56 percent. Note also that the pool of resume in our sample will consists of three main types of individuals: those out of job, on-the-job searchers and recent graduates.

To check for the representativeness of the job ads in our sample, we make use of the job vacancies reported in the 2013 Thailand Labour Demand of Establishment Survey (the LD thereafter). As noted, the LD only surveyed selected industries- which may not be the same industries that recruit using online tools. For the resume side, we cross-check our job portal sample with the Household Survey on Information and Communication Technology (ICT) of 2015 (ICT thereafter).

From Table 3.4, the ICT indicates that different age groups do not share the same online behaviours. Among those who use internet, the majority of individuals who use internet to help with job seeking activities concentrate among aged 20-29 (counted as 42% of those who use online tools for job-search). In the ICT, together, there are just over 1 million users who ever used internet for job searches. Our total count of resumes (total of 89,000 resumes updated within 2016) equals to 8.4% overall.

From Table 3.5-3.7, we find that, as mentioned in Khun and Shen (2013), our job ads over-represent higher skill candidates than the LD summary. Among different occupations, the comparison across different datasets indicates that our job portal database (micro-data and aggregate data) nearly mimics the exact number of total labour demand revealed in the LD survey for the Executives and the Professional occupations. However, in total, the sample of job ads from the internet falls significantly short of the total number of vacancies calculated in the LD. This is because the internet job portal extremely under-represent the demand for labour in low skilled occupations (namely agriculture, plant operators and services). We find this to be aligned with what suggest in the literature and that it is not a Thailand specific case.

### 3.4. Thailand's labour market tightness from aggregate-level job portal database

Here, we start by presenting the aggregate labour tightness (defined as the ratio of vacancies to resumes) at the 1-digit occupation level. (Figure 3.1) Using the data of the Job Portal No.1 from the first half of 2012 to 2016 (August 2016), we notice that the ranking of labour tightness remains somewhat persistent across the period- with the service occupation category has the highest v/u and the clerk category places at the bottom. The trend is declining, due mainly to the fact that the vacancies are on the rise with a much smaller rate that the rise of job seekers.

At the regional level, the ratio of v/u from the lowest is North, North-east and South. And this is consistent across seven years of data. For these regions, the size of vacancies had remained very low, in comparison to the number of job seekers applying with the region. While the (reverse) labour tightness looked relative high for East between the years 2011-2012, its level dropped down thereafter. Largely, this decline stems from a slower rate of vacancy expansion than the size of job seekers in the East region. Figures 3.2A-3.2C plot regional levels of total job ads, total resumes and the measure of labour tightness over time.

More interestingly, when we examine the spatial distribution of labour tightness (at province-level), we notice that there might be a crude correlation of higher vacancies for provinces with better road transport access. Figure 3.3A displays a heat-map of province-level labour tightness of 2016 and Figure 3.3B shows an overlay of locations of main railways in Thailand.

### 3.5. Learning about Thailand's labour market mismatch from micro-data job portal database

Now, we turn to our micro-data of a popular job portal. As explained in the previous section, we were able to access individual-level information from the site (Job Portal No.2). Without uncover any personal private identification, we extracted all job ads and all resumes (set for public accessibility) of the list of variables (shown in Table 3.2).

Using information at this level, we begin by investigating if the high level of labour market tightness in some occupations (found in the last section) might be due to mismatch of skill requirements of jobs and potential applicants. Using this dataset, we can able to sub-category the job ads further-by education, gender, or desired years of work experience. Figure 3.4 shows the geographical distribution of skills. For simplicity, we define applicants who obtain their qualification from technical vocation training, technically-focus university (Rajchamokol) and those with technical fields as those with hard degree. For those with commercial vocations or of other non-technical fields as with soft degree. Figure 3.4 compares the demand for applicants with bachelor or above from the job ads and the supply of job seekers with soft-degree. The comparison suggests that for most parts of the country, skills are located where they are demanded. However, when we compare Figure 3.4A and 3.4B to Figure 3.5C below, we notice that the provinces with high demand for college education are of those provinces with high market tightness (v/u)

In general, labour market friction is also believed to have cause by geographical mismatch of jobs available and applicants who are willing to relocate to a particular site. Note before that job ads in our database are heavily based in Bangkok (74% of total vacancies) whilst 61% of applicants are seeking jobs in Bangkok. Thus, we begin this examination by mapping the geographical distribution of jobs and separately of applicants across Thailand (based on their stated locational preference in the resume). Figure 3.5 shows that potential workforce do locate where most jobs are, with the exception of some lower North-eastern provinces. Nevertheless, the number of job seekers far outweighs total vacancies among popular locations. Therefore, when we plot the market tightness from our micro-data (Figure 3.4C), we observe that v/u is relatively low in these provinces. However, notice that along the eastern seaboard, our measure of market tightness looks relatively large.

Given that three quarter of total jobs are based in Bangkok, we look further if the geographical distribution of jobs and applicants, at the district level, display the same story within Bangkok. While the pool of Bangkok job seekers spread somewhat evenly across Bangkok 50 districts (with the exception of Ladprao, Bang Prapi, Bang-Na), job ads distribute much less uniformly. Bang Rak and Wattana have the highest (8% share) of total Bangkok job vacancies. Figure 3.6A-C displays the geographical spread of vacancies, applicants and market tightness around Bangkok. We can observe that within Bangkok, unfilled jobs locate either in the very central of the capital or around the outer districts. In contrast, preferred locations of work, indicated by applicants, concentrate more in the northern parts of the capital.

We list out two potential factors, which may contribute to the geographical mismatch of this selected Thai labour market. First, the presence of geographical immobility (that is job seeker's unwillingness to relocate) may account for unfilled vacancies in certain provinces. Second, within a province, a gap between the minimum accepted salary (the so-called reservation wage) and the maximum pays (the reservation productivity of workers) can potentially reduce the matching of applicants into jobs.

To support the first explanation, we take the location of the institution where the latest qualification was obtained as our proxy for the location of most current residence. We then map this location of residence to the location of work. Figure 3.7 presents the share of individual applicants' resident locations within each location choice of work. Whilst Bangkok can attract nearly 50% of applicants from outside its region, most provinces in the North and the Northeast have failed to reach 10%. This also means that the majority of applicants in the North and the Northeast are drawn from those who had studied in the regions.

To see if other observable individual characteristics may influence a choice of location for work, we run a regression analysis using our resume database. Let $Loc_i$ be an indicator with value of 1 if the individual chooses her work location the same as her residence, $\boldsymbol{X}_i$ be a vector of individual characteristics observed in the dataset. Our mobility regression is as below. We run a province-of-work fixed-effect model to account for unobservable variables at this level. The standard errors are also clustered at the region level.

$$Loc_i = \boldsymbol{X}_i\boldsymbol{\gamma} + \mu_i \qquad \text{(Equation 7)}$$

We can see that being a female job seeker is associated to 0.2% more likely to stay in the same area as the last education. The higher the years of work experience, the more likely a job seeker will stay put. Young job seekers are more geographical mobile. On average, individuals with higher grade point average are more likely to move away from where they previously studied. The regression results are in Table 3.8.

Another source of labour mismatch can be due to the gap between the reservation wage and the expected labour productivity. Under a matching model, a worker would always reject the job offer if the offered wage is not at least equal to her reservation wage. Likewise, upon the initial match with a worker, the match will break if the worker is not as productive as the offered wage by the firm (Mortensen and Pissarides, 1994). Therefore, to investigate this potential mechanism, we look at salary differences within a well-defined job (matched job-applicant with the same level of qualification, experience and occupation demanded and supplied). Table 3.9 displays the average salary for each of these jobs.

We also run a basic Mincerian-style equation, separately for the demand side (job ads database) and the supply side (resume database). We do this to check if the applicants and firms share similarities in their wage-setting behaviour. In other words, we want to see if each additional skill or characteristics share the comparable market value for each side of the market.

Let $W_j$ ($W_i$) be the log of level monthly-rate salary offered (requested) by a job ad $j$ (an applicant $i$), $\boldsymbol{Z}_j$ ($\boldsymbol{Z}_i$) be a vector of individual job (applicant) characteristics observed in the dataset. Control variables include occupation categories and locations. The standard errors are also clustered at the region level. We derive Equation 8 for the job ads and the wage offer. Equation 9 represents the reservation wage equation, for the applicant side.

$$W_j = \boldsymbol{Z}_j \boldsymbol{\pi_j} + \varphi_j \qquad \text{(Equation 8)}$$

$$W_i = \boldsymbol{Z}_i \boldsymbol{\pi_i} + \rho_i \qquad \text{(Equation 9)}$$

We run a simple OLS Mincerian-style wage equation, each for the demand and the supply side of this particular labour market platform. By using log-form of monthly salary, all coefficients can now be interpreted as percentage points. Table 3.10A displays the results with log of wage offer (at job ad level) as the dependent variable. Table 3.10B presents the results with log of wage request (at individual resume level) as the dependent variable. In both equations, we controls for the same set of 1-digit occupation categories and province-of-job fixed effect. In both regressions, the standard errors are robust and clustered at the regional level.

From the wage offer side, it reveals that (under the full specification- Model III) job ads which target female only is related to just over 10 percent reduction in the wage offer. This penalty is not found for male-only job ads. Jobs seeking young workers look to offer smaller wage (18% negative for jobs targeting adolescents and 8% for jobs for aged 26-35). Work experience is valued by the demand side. One additional year of work experience is related to 10% increase in the wage offer. Jobs that do not specify a level of education qualification pays 6% more than jobs seeking bachelor graduate (bachelor degree is our omitted education category here). While bachelor degree or highest tops the salary level, jobs looking for the minimum education of Matayom 3 (at present, the school-leaving minimum qualification) pays the least. Graduating with Matayom 6 is valued less than graduating with a High Vocational School qualification in the job ads database.

Looking now at the reservation wage side, strikingly female applicants tend to ask for nearly 10% less wage than the males. Young job seekers (under 25 years) tend to set higher reservation wage than slightly older cohorts. However, the most senior age group is found to set the highest level of reservation wage in this database. Consistent with employers, job seekers do relate their work experience to higher requested level of pays. Without controlling for past salary (specification I), higher years of reported work experience is related to higher asking pay. However, once past salary is added, the magnitude of years of experience reduces around 50 percent. But overall, we observe a non-linear increase in the rise of experience and the rise of asking pays. In comparing to applicants seeking full-time jobs, those who look for freelance do not reduce their asking wage. In contrast, those looking for internship or part-time job ask for roughly 25% less pay.

Interestingly, applicants value their school performance (proxy by Grade Point Average) as well as their self-report skills (English competency and other languages). With applicants of general bachelor degree as the education baseline, those with vocational qualifications (fewer years of schooling) request less pays. Notice that those of commercial VQ demand almost 3 times less pays than those of technical VQ. Among bachelor graduates, applicants with a Rajapat degree is correlated with 7% lower wage reservation than the general degree. However, those with a Rajamonkol on average set 6% higher wage.

### 3.6. Is mismatch as a result of gender-targeting job opening?

Our argument is that because of many jobs tend to limit at the outlook, the gender of workers (be that discrimination or statistic learning), jobs are not filled and some jobs are over-subscribed. This, of course, causes disadvantages to the firms: unfilled vacancies are a productivity loss. And filling job with under-qualified but right gender also leads to productivity loss. For workers, it is a gain for those who not subjected to discrimination but loss to those who are. Furthermore, this affects the occupation paths and planning of future labour forces. Thus, national output is a risk.

From the dataset, we investigate on-the-surface, if this feature of gender-mismatch may take place in this sub-labour market. We begin by investigating the nature of gender-targeting job ads. By analysing a Chinese internet-based job portal, Kuhn and Shen (2013) find a *negative* correlation between the probability to gender-target their ads and the level of skills required in a job (measured by qualification level, work experience or occupational ranking). This suggests that the firm’s strategy to gender-target goes hand in hand with its aim to skill-target.

From Table 3.11, it is plain to see that more than 30 percent of our sample of job ads (in the micro-data) is gender-target[24]. To investigate further, we group these job ads into different occupational categories, educational requirement, minimum work experience and age-group of ideal candidates. Figure 3.8 plots the share of job ads by its gender target and 1-digit occupation code. By ranking from most highly skilled occupation (executives) to least skilled occupation (services), we observe that the share of gender-neutral job ads gets smaller as the skill requirement declines. We check with the 2013 Labour Demand Survey to see if the corresponding stylise is found. Figure 3.9 reflects a consistent story between our two samples.

More strikingly, as education requirement gets higher (ranking from minimum requirement of lower secondary school to master degree), we notice a rise of gender-neutral job ads (Figure 3.10). Again, we find the similar but weaker correlation when we plot the share of gender-target job ads with work experience.

However, when we examine the gender-skill targeting relationship using age requirement as a proxy (as age is generally translated to experience), we find the contrary (Figure 3.11). Job ads that aim at young candidates (aged 15-19) are less gender-specific that job ads for those of 25 years and above. Instead, among those gender-target job ads that aim for older candidates, a higher proportion of jobs target male candidates more heavily (from 23 percent for aged 20-24 to 27 percent for aged 30 or above). These leave much room for interpretation. Among the pool of younger candidates, both males and female are equally less likely to have a family. As the pool gets older, females become less ideal workers to a firm. Hence, job openings which aim to recruit older workers are more likely to target and screen for males.

## 4. Summary and conclusions

In the era of the digital economy, technology has become a part of our everyday lives. The way we do our economic activities has changed from how we did them in the traditional economy. The labour market in the digital age is also different from the traditional labour market. Specifically, both employers and job applicants rely more on the internet when posting

[24] Lawler and Bae (1998) looked at a sample of 902 job ads by MNCs in Thailand from the Bangkok Post (English language newspaper) and checked for gender-targeting behaviour of big corporations. They found that 25.2% were male-only, 13.9% were female only and 60.3% were gender-non-specific. Among only Thai firms, 32.5% were male-only, 13.1 were female only.

vacancies, applying for jobs, and searching for information relating to employment benefits. These internet activities create the digital footprints that can be used to analyse the labour market in the way that could not be done by utilising the traditional labour data.

The objective of this study is to use the internet-based data (e.g. Google Trends data, and the internet job board data) to analyse the labour market from a new perspective. Our study demonstrates that: (i) the data from Google Trends could be used to 'nowcast' the labour market indicators (namely, the unemployment rate and the registration for the unemployment). (ii) The data from the internet job boards allow us to understand the labour market from both the demand side (firms' vacancy postings) and the supply side (applicants' resume postings). In addition, we can also observe the distribution of jobs and workers across space and study the spatial concentration of jobs and workers of specific skills and occupations. (iii) This firm-applicant dataset collected from the internet job boards also allow us to study the disparity between the offered wage (by firms) and the asking wage (by applicants) across types of jobs, locations, and skills. (iv) And with careful analysis of the internet job board data, we can extract a proxy for labour market tightness at the cross section of job characteristics, across space, and over time.

The use of internet-based data to analyse the labour market may still be uncommon in Thailand. And the access to such data may still be limited. However, we believe that these data provide new meaningful insights for the Thai labour market. Therefore, it could be helpful for the policymakers to pay attention to these data along with the traditional labour data when making policy decisions.

## References

Antenucci, D., Cafarella, M., Levenstein, M. C., Re, C., & Shapiro, M. D. (2014, March). Using Social Media to Measure Labor Market Flows. NBER Working Paper No. 20010.

Askitas, N., & Zimmermann, K. F. (2009, June). Google Econometrics and Unemployment Forecasting. IZA Discussion Paper Series No. 4201.

Autor, D. H. (2001). Wiring the labor market. *The Journal of Economic Perspectives*, *15*(1), 25-40.

Barnichon, R. (2010). Building a composite help-wanted index. Economics Letters, 109(3), 175-178.

Barnichon, R., & Figura, A. (2015). Labor market heterogeneity and the aggregate matching function. *American Economic Journal: Macroeconomics*, *7*(4), 222-249.

Bell, V., Co, L. W., Stone, S., & Wallis, G. (2014, Q1). Nowcasting UK GDP Growth. Bank of England Quarterly Bulletin.

Bragoli, D., & Modugno, M. (2016). A Nowcasting Model for Canada: Do U.S. Variables Matter". Board of Governors of the Federal Reserve Finance and Economics Discussion Series 2016-036.

Carriere-Swallow, Y., & Labbe, F. (2013). Nowcasting with Google Trends in an Emerging Market. Journal of Forecasting, 32, 289-298.

Choi, C., & Varian, H. (2009a, April). Predicting the Present with Google Trends. Google Inc.

Choi, C., & Varian, H. (2009b, July). Predicting Initial Claims for Unemployment Benefits. Google Inc.

Diamond, P. A., & Şahin, A. (2015). Shifts in the Beveridge curve. *Research in Economics*, *69*(1), 18-25.

Edelman, B (2012) "Using internet data for economic research" J Economic Perspectives

Fabo and Tijdens (2014) "Using web data to measure the demand for skills" CELSI Discussion paper no. 21.

Fonduer, Y., & Karame, F. (2013). Can Google Data Help Predict French Yough Unemployment? Economic Modelling, 30, 117-125.

Giannone, D., Reichlin, L., & Small, D. (2005). Nowcasting GDP and Inflation: The Real-Time Informational Content of Macroeconomic Data Releases. Board of Governors of the Federal Reserve Finance and Economics Discussion Series 2005.

Hara, N., & Yamane, S. (2013, October). New Monthly Estimation Approach for Nowcasting GDP Growth: The Cast of Japan. Bank of Japan Working Paper Series No. 13-E-14.

Higgins, P. (2014, July). GDPNow: A Model for GDP "Nowcasting". Federal Reserve Bank of Atlanta Working Paper Series 2014-7.

Khun, P., & Shen, K. (2013). Gender Discrimination in Job Ads: Evidence from China. Quarterly Journal of Economics, 128(1), 287-336.

Knotek II, E. S., & Zaman, S. (2014). Nowcasting U.S. Headline and Core Inflation. Federal Reserve Bank of Cleveland Working Paper 14-03.

Krueger (2000) "The internet is lowering the cost of advertising and searching for jobs" New York Times, July 20th.

Kuhn and Mansour (2014) "Is internet job search still ineffective" EJ

Kuhn and Shen (2013) "Gender discrimination in Job Ads: Evidence from China" QJE

Kuhn and Skuterud (2004) "Internet job search and unemployment durations" AER.

Kurekova, L. (2014) "Using Internet Data to Analyse the Labour Market: A Methodological Enquiry" (IZA)

Lawler, J. J., & Bae, J. (1998). Overt employment discrimination by multinational firms: Cultural and economic influences in a developing country.*Industrial Relations: A Journal of Economy and Society*, *37*(2), 126-152.

Lenaerts, K., Beblavy, M., & Fabo, B. (2016). Prospects for Utilisation of Non-Vacancy Internet Data in Labour Market Analysis - An Overview. IZA Journal of Labor Economics, 5(1).

Luciani, M., Pundit, M., Ramayandi, A., & Giovanni, V. (2015). Nowcasting Indonesia. Board of Governors of the Federal Reserve Finance and Economics Discussion Series 2015-100.

Marinescu, I. (2016, July). The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board. NBER Working Paper No. 22447.

Modugno, M., Soybilgen, B., & Yazgan, E. (2016). Nowcasting Turkish GDP and News Decomposition. Board of Governors of the Federal Reserve Finance and Economics Discussion Series 2016-044.

Mortensen, D. T., & Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, *61*(3), 397-415.

Nakamura et al (2010) "Jobs Online"

Sahin, A., Song, J., Topa, G., & Violante, G. L. (2011) Measuring Mismatch in the US Labor Market. mimeo

Sahin, Aysegul, Joseph Song, Giorgio Topa, and Giovanni L. Violante. "Measuring mismatch in the US labor market." Manuscript, revised March (2011).

Seabold, S., & Coppola, A. (2015, August). Nowcasting Prices Using Google Trends: An Application to Central America. World Bank Group Policy Research Working Paper 7398.

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

Stevenson, B. (2009). The Internet and Job Search. In *Studies of Labor Market Intermediation* (pp. 67-86). University of Chicago Press.

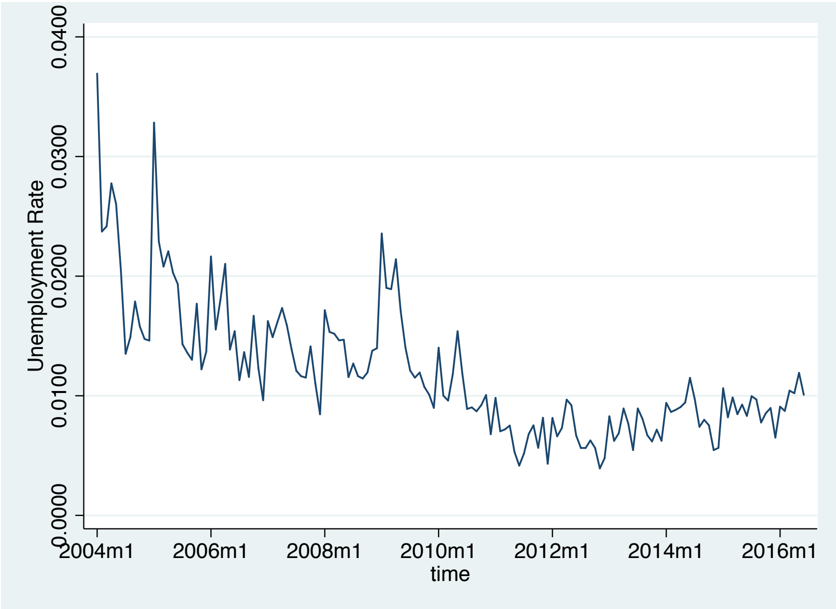
Suhoy, T. (2009, July). Query Indices and a 2008 Downturn: Israeli Data. Bank of Israel Discussion Paper No. 2009.06.

Vincente, M. R., Lopez-Menendez, A. J., & Perez, R. (2015). Forecasting Unemployment with Internet Search Data: Does It Help to Improve Predictions when Job Destruction is Skyrocketing? Technological Forecasting & Social Change, 92, 132-139.

https://www.conference-board.org/data/helpwantedonline.cfm

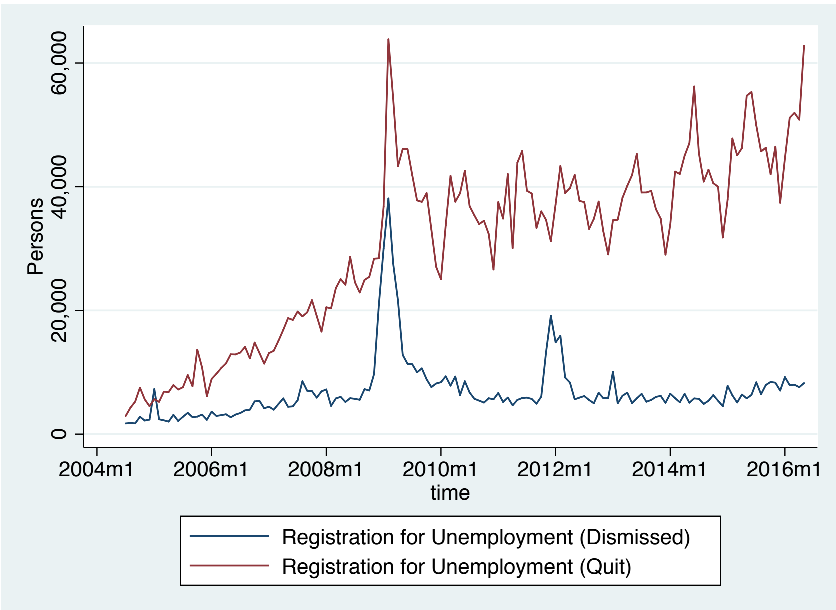
**Figures**

Figure 1.1: Thailand's unemployment rate from 2004 to early 2016



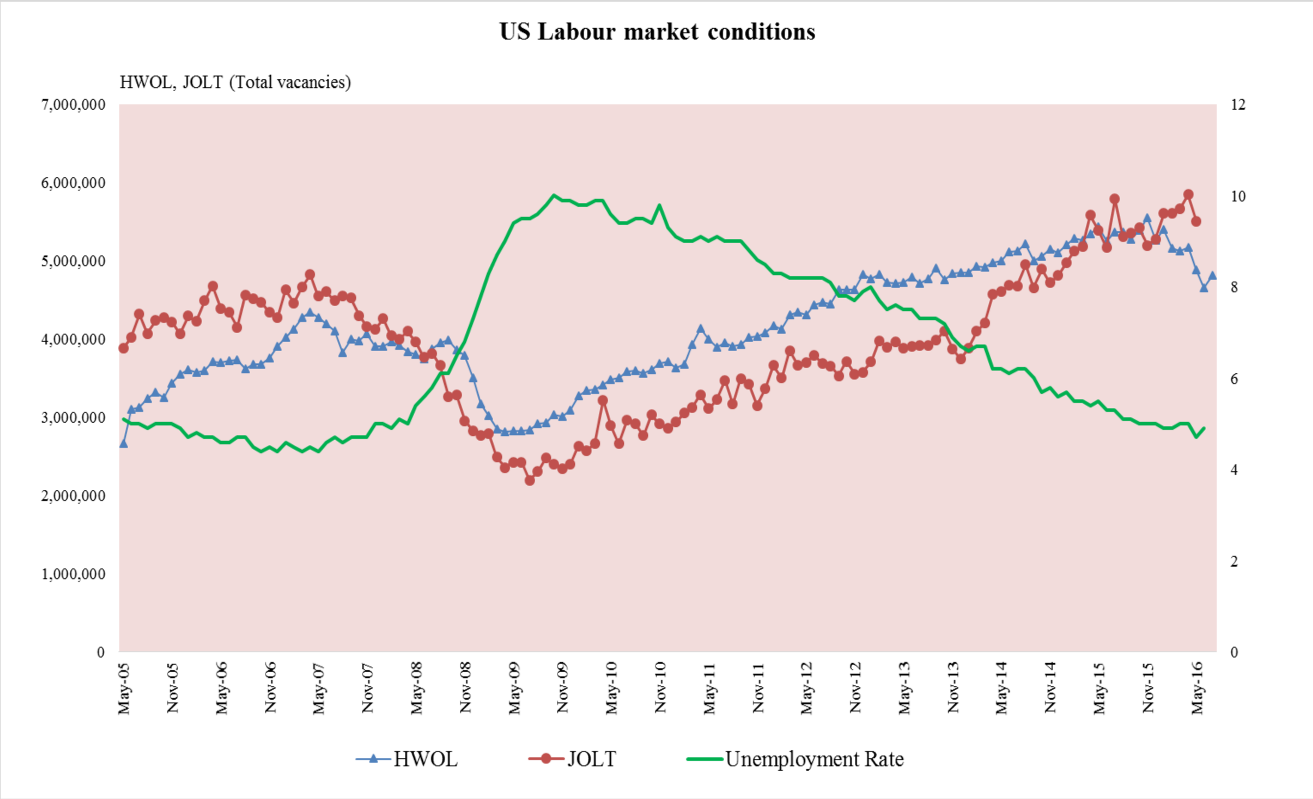

Source: Labour Force Survey, National Statistical of Thailand, 2004-2016

Figure 1.2: Thailand's Registration for Unemployment



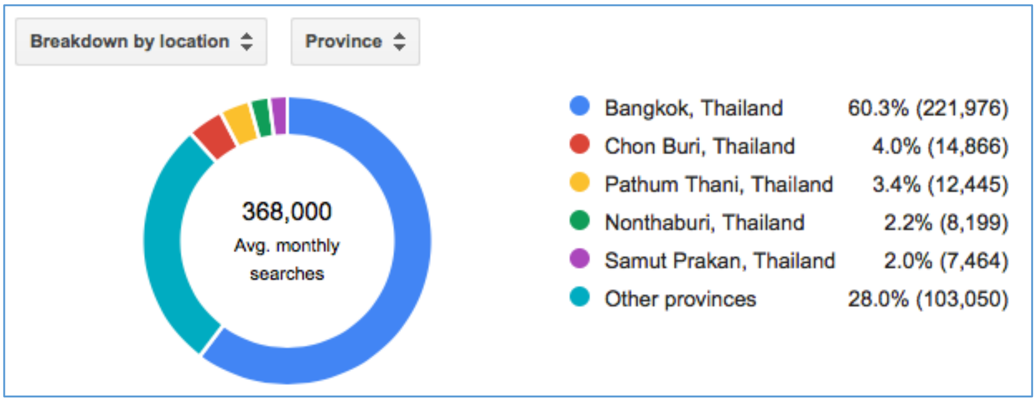

Source: Department of Employment, Ministry of Labour, 2004-2016

Figure 1.3: Trends of US unemployment rate (US CPS) and the job vacancies (JOLTS and HWOL)



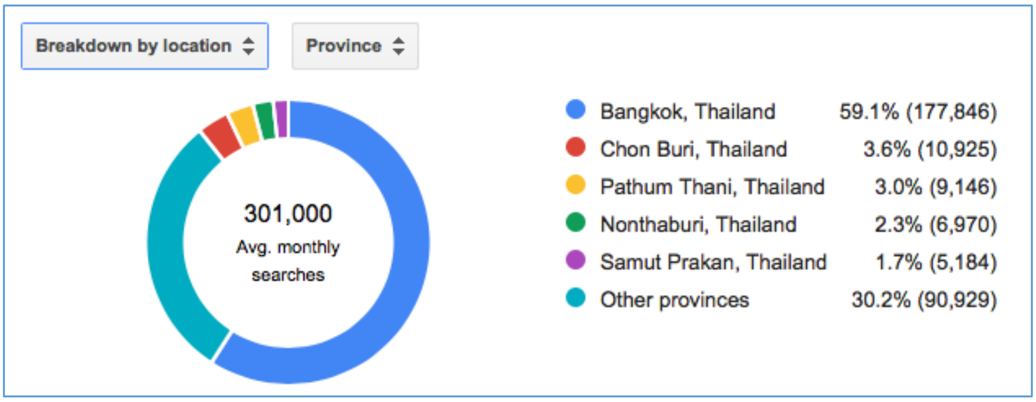

Sources: BLS, the Conference Board

Figure 2.1: Average Monthly Search for "หางาน" (Job Search) Breakdown by Location



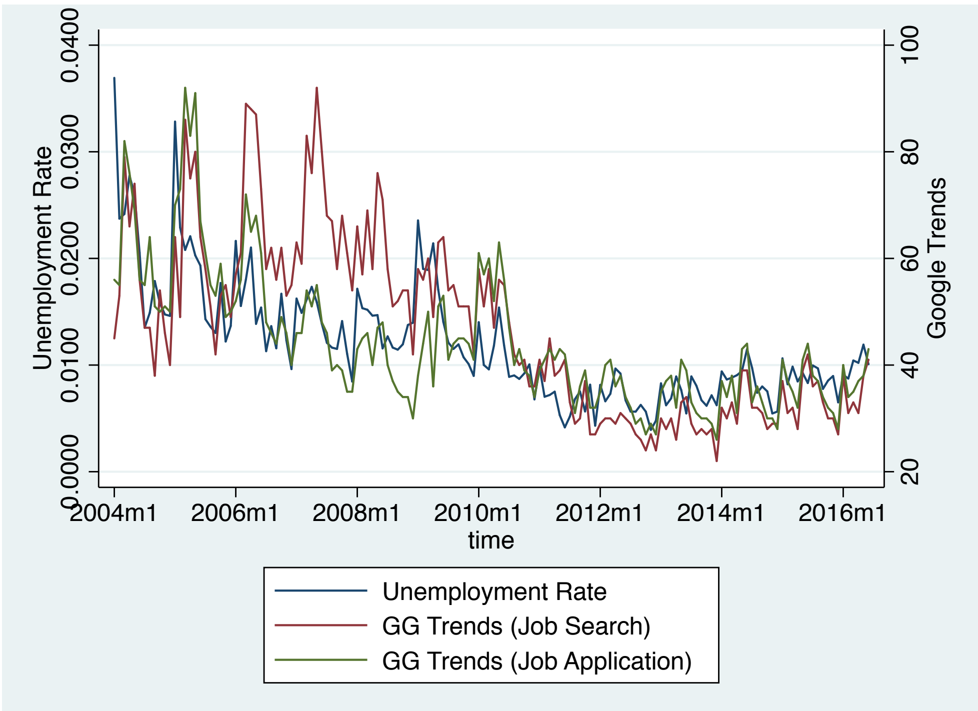

Source: Google AdWords' Keyword Planner

Figure 2.2: Average Monthly Search for “สมัครงาน” (Job Application) Breakdown by Location



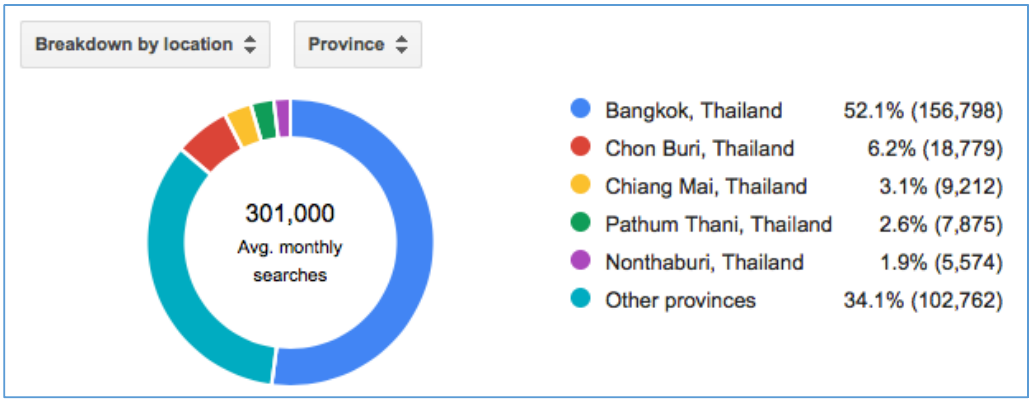

Source: Google AdWords' Keyword Planner

Figure 2.3: Thailand's Unemployment Rate with Google Trends



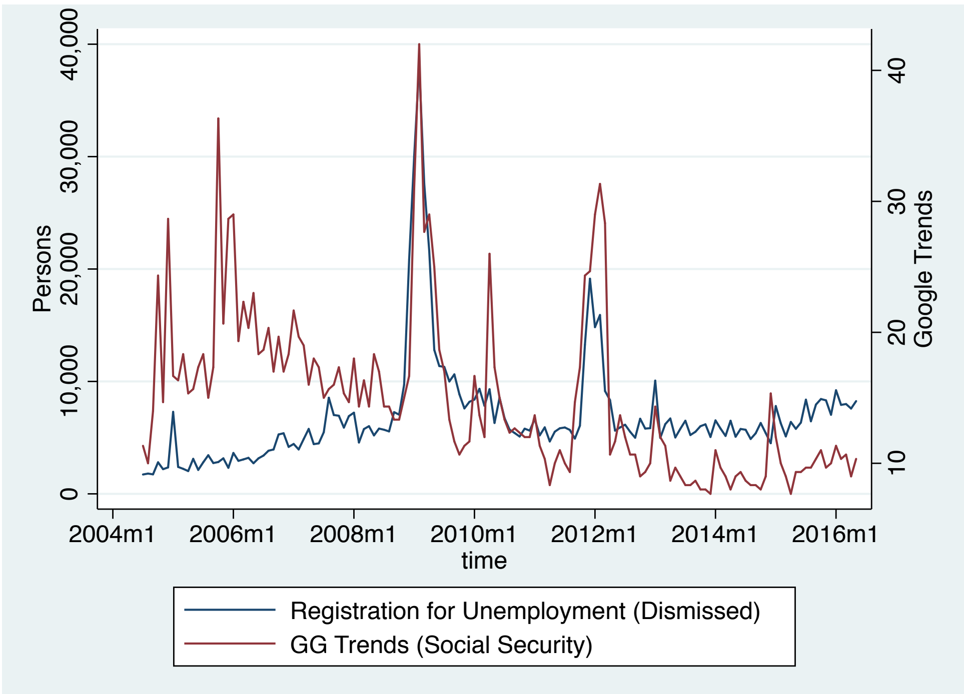

Source: Labour Force Survey, National Statistical of Thailand, 2004-2016; and Google Trends

Figure 2.4: Average Monthly Search for "ประกันสังคม" (Social Security) Breakdown by Location



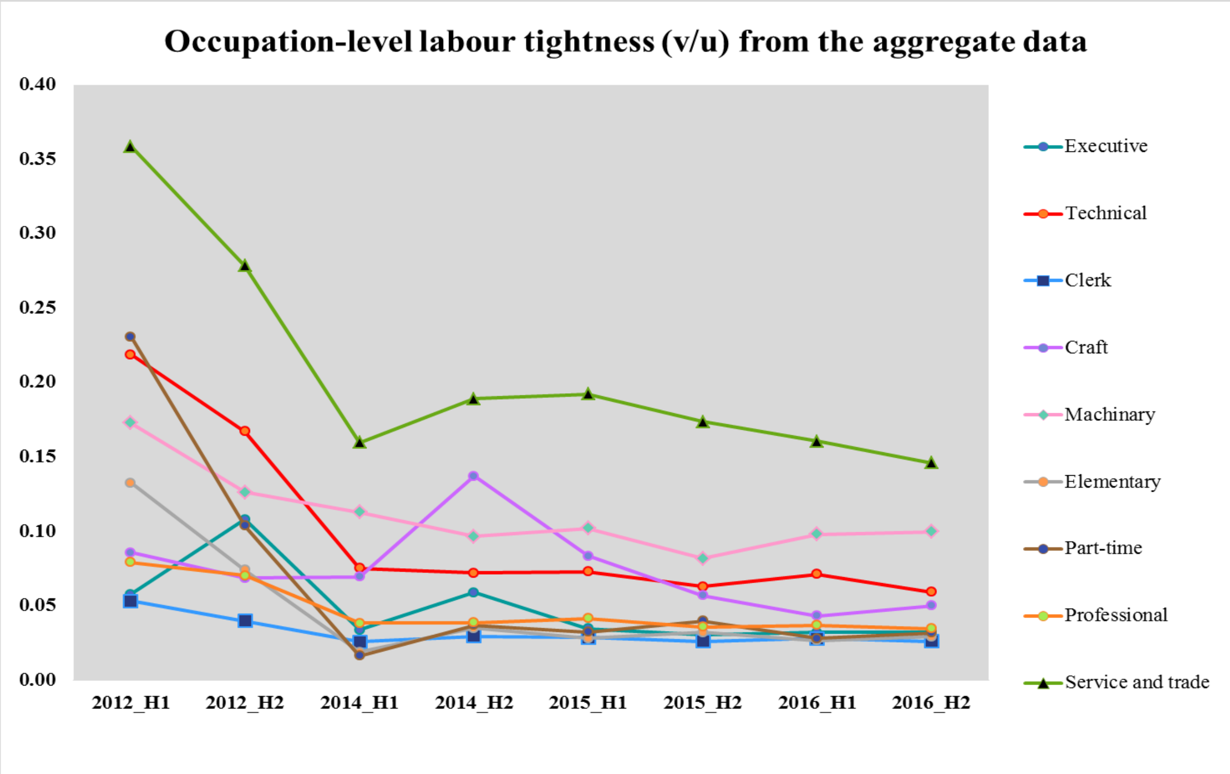

Source: Google AdWords' Keyword Planner

Figure 2.5: Thailand's Registration for Unemployment (Dismissed Workers) with Google Trends



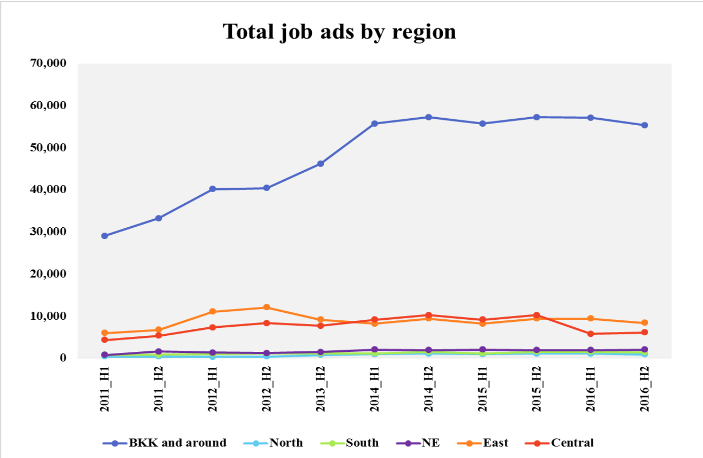

Source: Department of Employment, Ministry of Labour, 2004-2016; and Google Trends

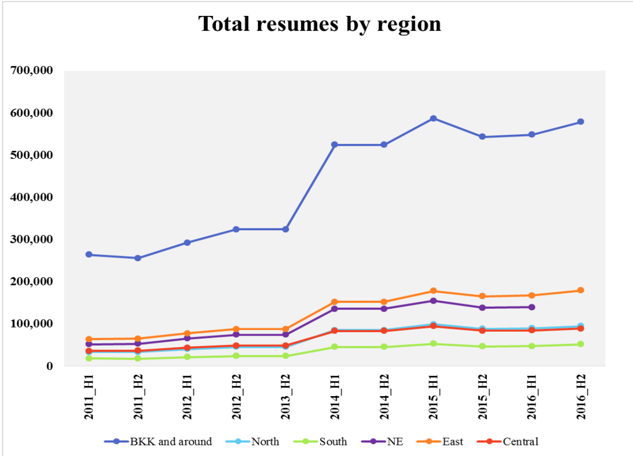
Figure 3.1: Trends of labour tightness by occupation

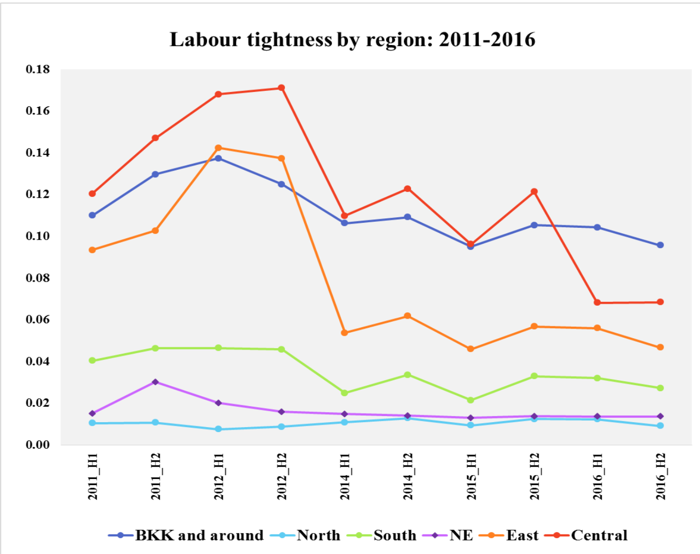



Source: Archived data of the Job Portal No.1

Figure 3.2A-3.2C: Total job ads, total resume and the measure of labour tightness over time by region

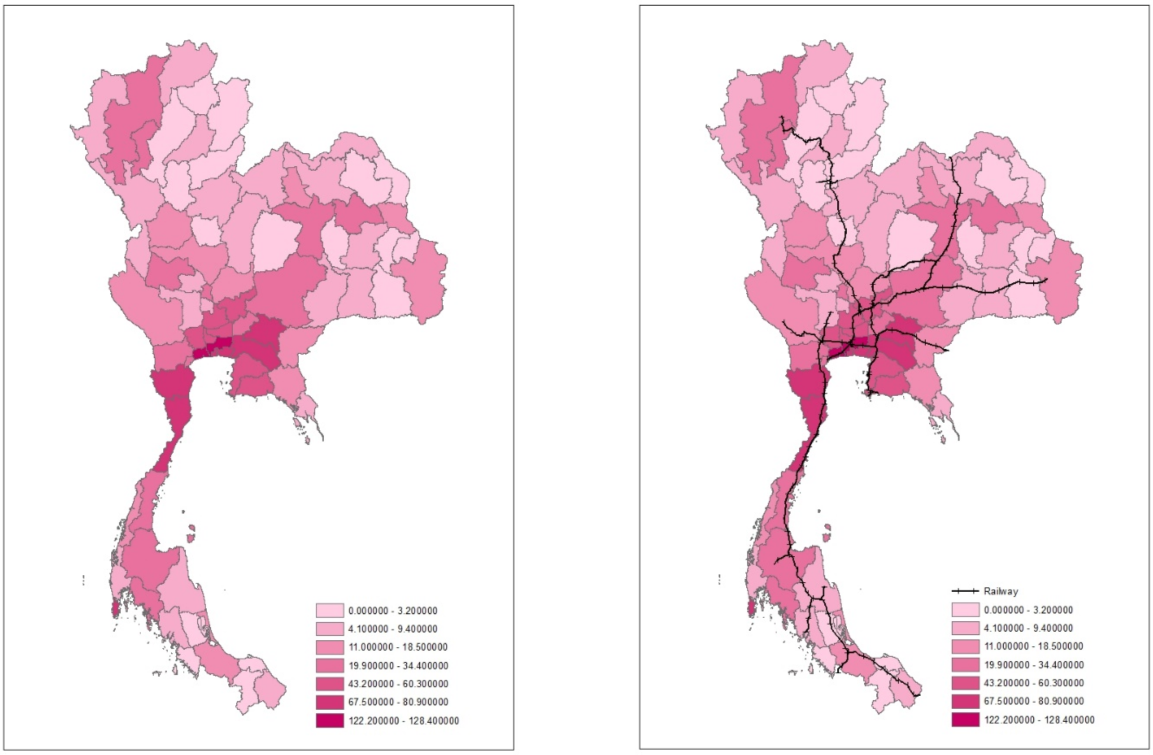



Source: Archived data of the Job Portal No.1

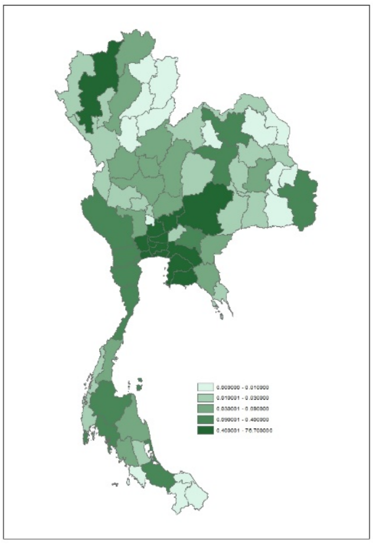
Figure 3.3A-B: Geographical distribution of labour market tightness with railways overlaid (aggregate data)



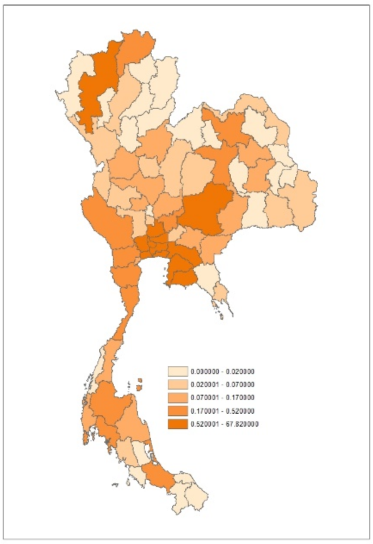

Sources: Archived data of the Job Portal No.1 (2016) and Ministry of Transports (2010)

Figure 3.4A-E: Distribution of demand for bachelor degree or above, compare to the supply of job seekers with the same qualification (as percentage of total jobs or resumes within each province)

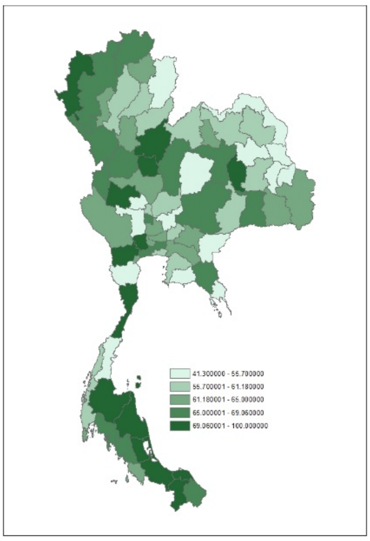
(A) Demand for bachelor degree



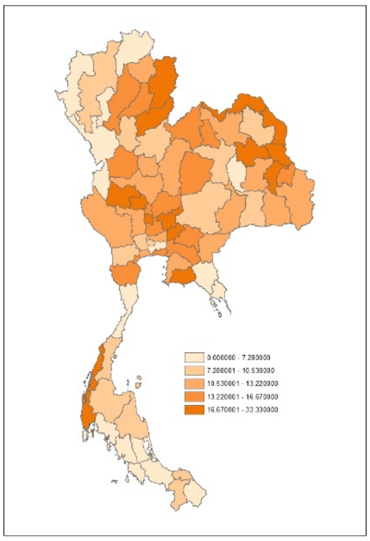

(B) Demand for high vocational degree



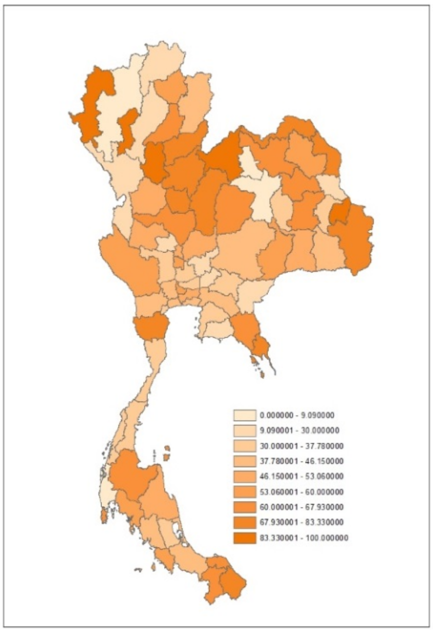

(C) Supply of bachelor degree




(D) Supply of high vocational degree



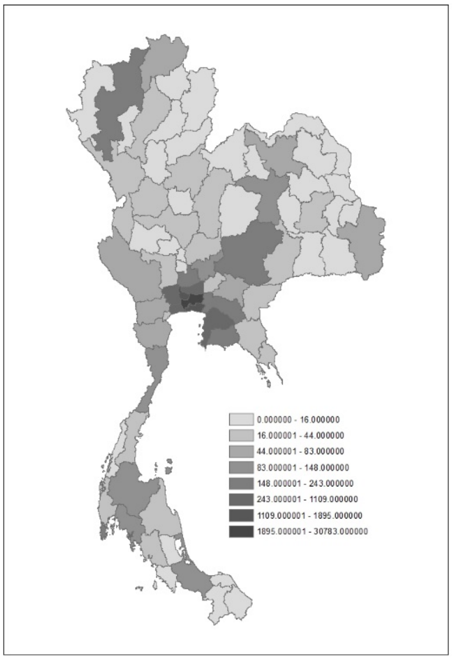
(E) Supply of soft-degree (% of total)



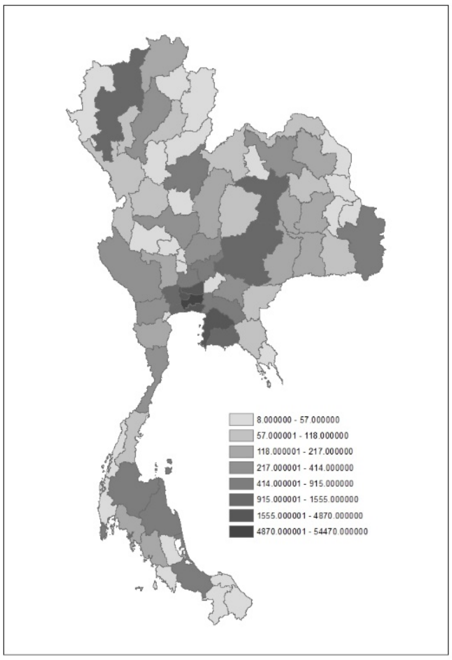

Source: The Job Portal No.2 (2016)

Figure 3.5A-C: Distribution of vacancies, resumes and market tightness around Thailand provinces (micro-data)

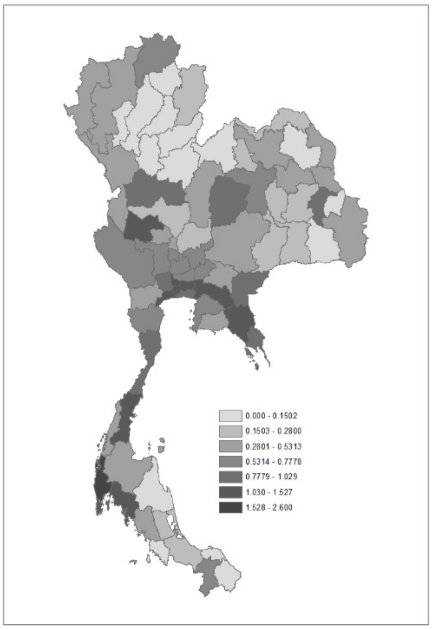
(A) Distribution of total vacancies




(B) Distribution of total resumes



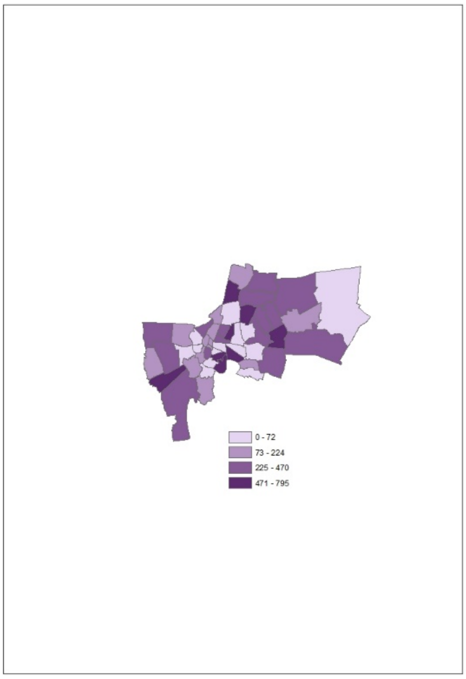

(C) Distribution of market tightness



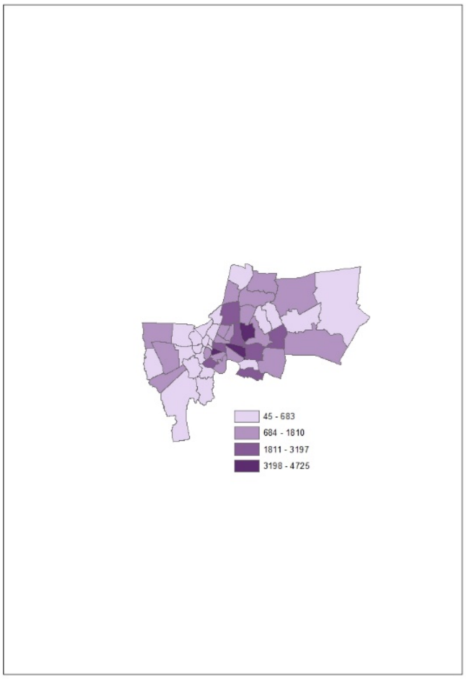

Source: The Job Portal No.2 (2016)

Figure 3.6A-C: Distribution of vacancies, resumes and market tightness (v/u) around Bangkok districts

(A) Share of total vacancies around Bangkok




(B) Share of total resumes around Bangkok




(C) Distribution of market tightness

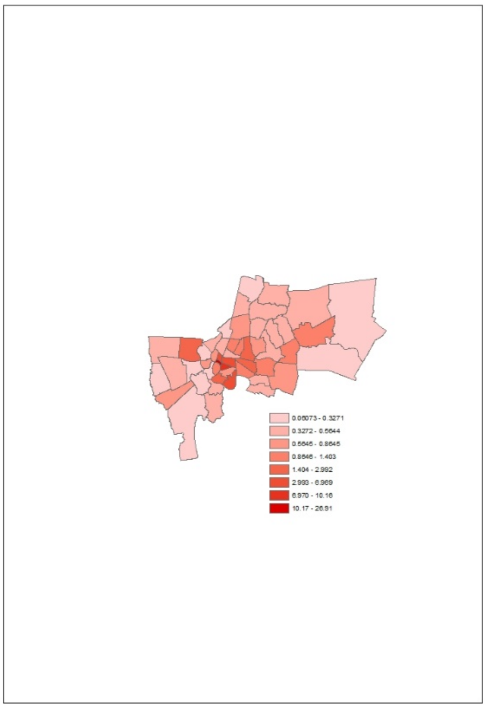
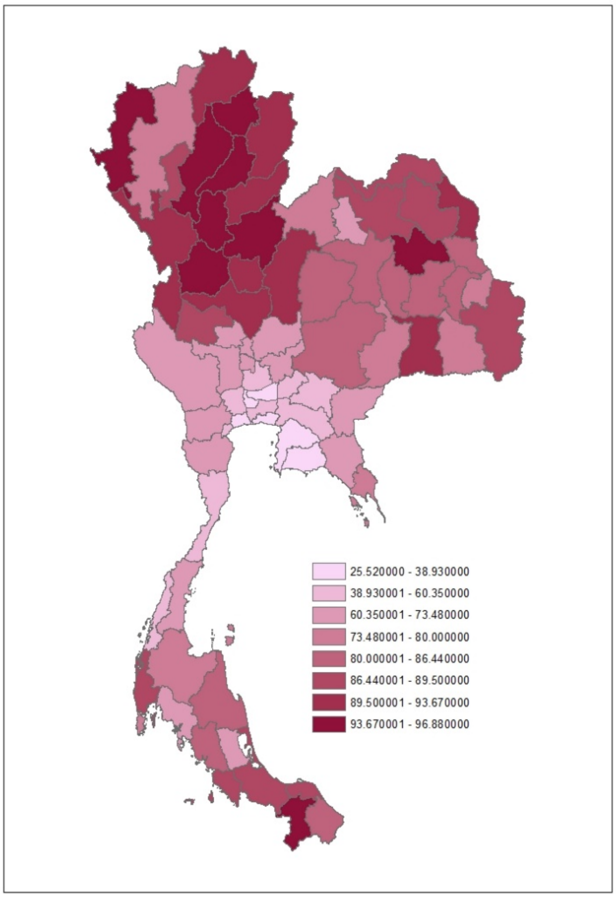
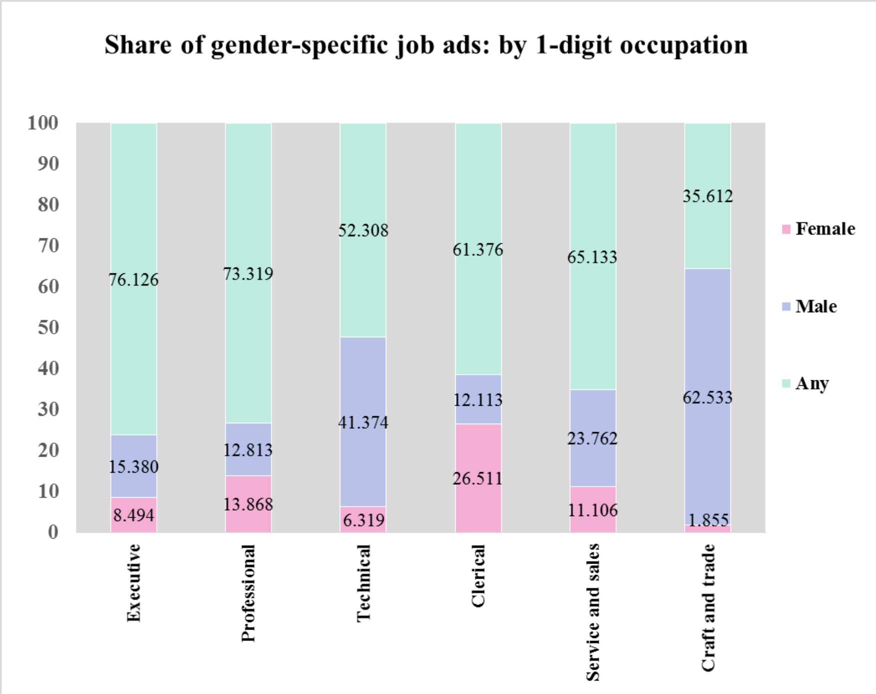


Figure 3.7: Share of applicants from different education location in a chosen province of work



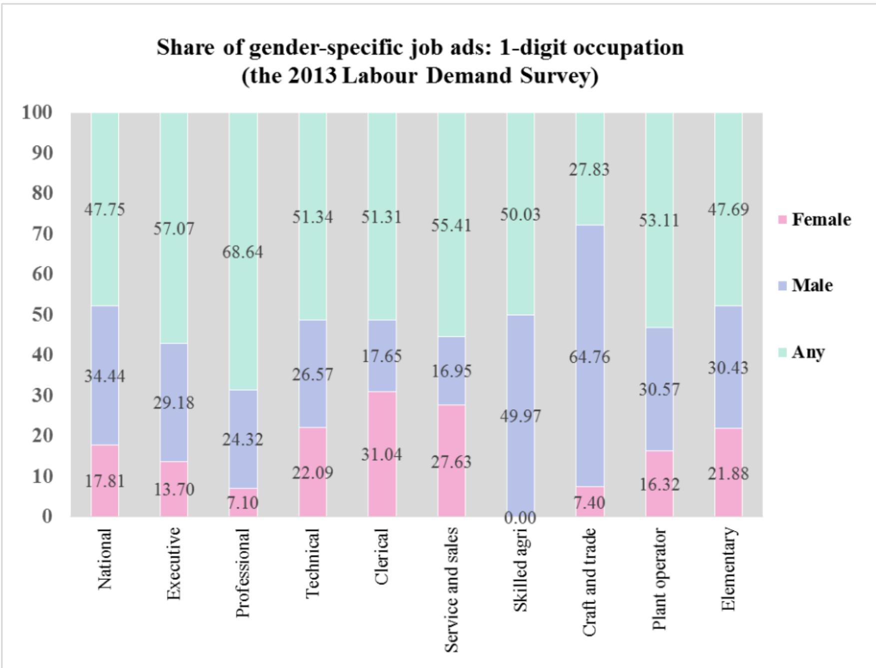

Source: the Job Portal No.2 (2016). The original region is defined as the region of the institution where an applicant obtained the latest qualification

Figure 3.8: The share of job ads by its gender target and 1-digit occupation code.



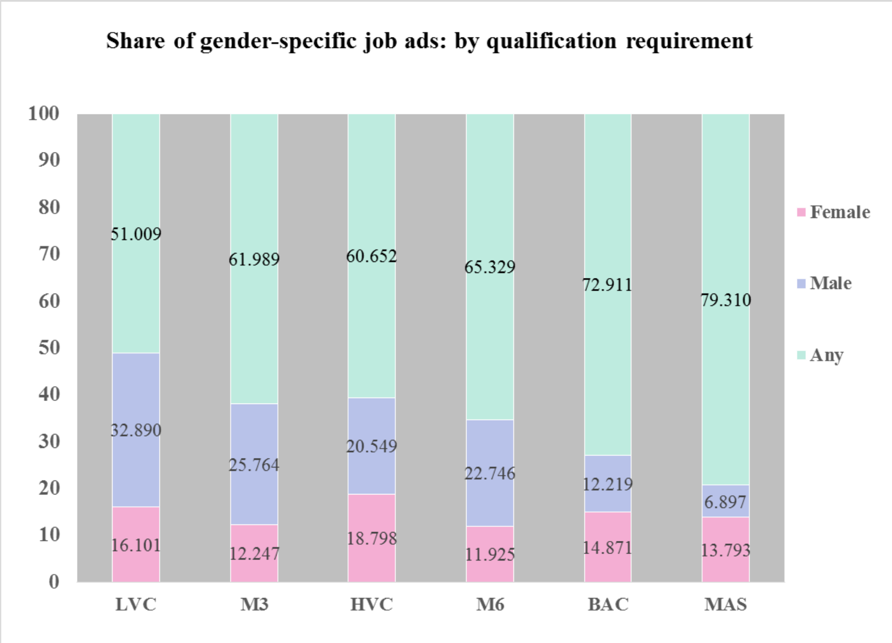

Source: The Job Portal No.2 (2016).

Figure 3.9: The share of job ads by its gender target and 1-digit occupation code from the LD survey

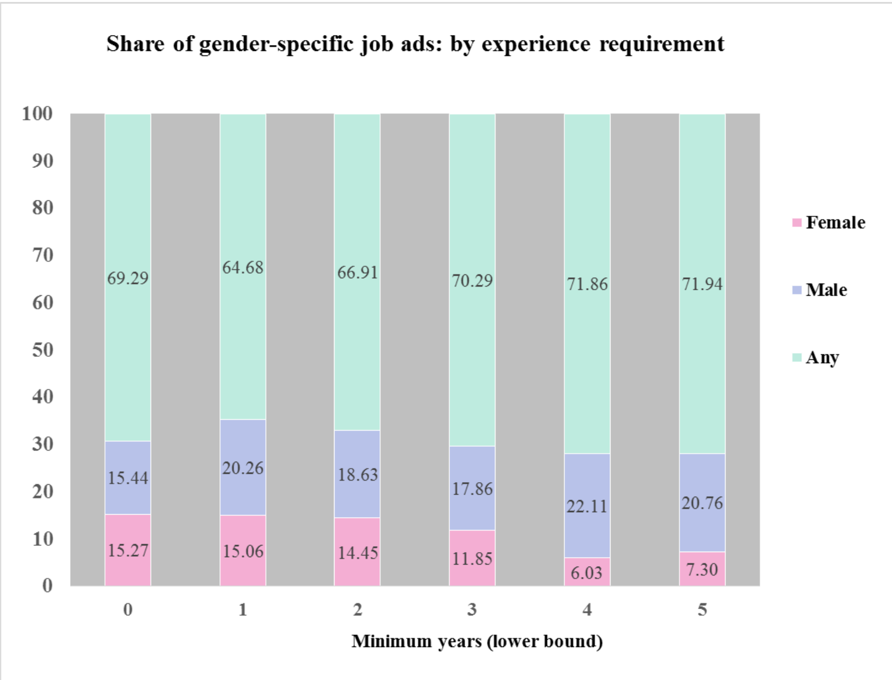



Source: The Labour Demand of Establishment Survey (2013).

Figure 3.10A-B: The share of job ads by its gender target and skills requirements (qualification and work experience)

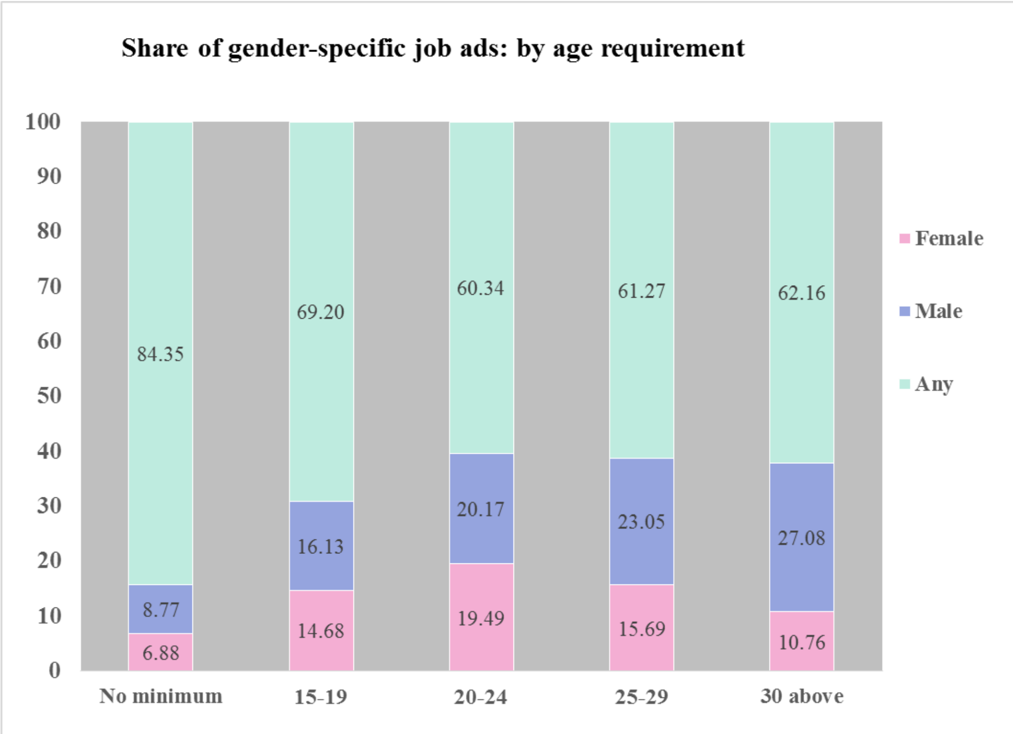



Source: The Job Portal No.2 (2016).




Source: The Job Portal No.2 (2016).

Figure 3.11: The share of job ads by its gender target and age requirements



## Tables

Table 1.1: Share of job search methods from the ICT Survey 2015

| | Total | Penetration rate of internet as job search platform |
|---|---|---|
| Looking for job or sending/submitting job application (aged 15 and older) | 1,062,208 | (%) |
| Total of employed (LFS 2016 Q1) | 37,000,000 | 2.870 |
| Total working age | 5,500,000 | 19.313 |
| Total unemployed (LFS 2016 Q1) | 370,000 | 287.083 |
| Total under-employed (LFS 2016 Q1) | 237,500 | 447.245 |
| Total under-employ & unemp (LFS2016 Q1) | 607,500 | 174.849 |

Source: LFS 2016 Q1 and ICT Survey 2015

Table 2.1: Variables used for nowcasting unemployment rate

| Variables | Frequency | Release Date |
|---|---|---|
| Unemployment Rate | Monthly | 5th Day of the Following Month |
| Google Trends หางาน (Job Search) | Weekly | Real-time |
| Google Trends สมัครงาน (Job Application) | Weekly | Real-time |

Table 2.2: Unemployment regression results

| | I | II | III | IV |
|---|---|---|---|---|
| A | 0.00278*** | 0.00043 | -0.00074 | -0.00111 |
| | (.00062) | -0.00056 | -0.00104 | -0.00101 |
| b1 | 0.753*** | 0.539*** | 0.51064*** | 0.451*** |
| | (.0560) | (.0512) | (.0736) | (.065) |
| c1 | | 0.0001*** | | 0.00007*** |
| | | (.000019) | | (.000021) |
| c2 | | | 0.00015*** | 0.0001** |
| | | | (.00004) | (.000045) |
| R2 | 0.6535 | 0.7143 | 0.7147 | 0.7342 |
| Adjusted R2 | 0.6511 | 0.7104 | 0.7108 | 0.7287 |
| MSE | 0.000009 | 0.000008 | 0.000008 | 0.000007 |
| MAE | 0.002124 | 0.001873 | 0.001886 | 0.001805 |
| MAPE | 0.1925 | 0.1679 | 0.1729 | 0.1635 |

Notes: *** 1%, **5%, * 10%. Robust standard errors in parentheses.

Table 2.3: Variables used for nowcasting registration for unemployment (dismissed workers)

| Variables | Frequency | Release Date |
|---|---|---|
| Registration for Unemployment (Dismissed Workers) | Monthly | 5th Day of the Following Month |
| Google Trends ประกันสังคม (Social Security) | Weekly | Real-time |

Table 2.4: Registration for unemployment (dismissed workers) regression results

| | **V** | **VI** |
|---|---|---|
| A | 968.41 | 77.62 |
| | -660.74 | -1022 |
| b1 | 0.865*** | 0.826*** |
| | (0.113) | (0.095) |
| c1 | | 77.18* |
| | | (46.64) |
| R2 | 0.7551 | 0.7635 |
| Adjusted R2 | 0.7533 | 0.7601 |
| MSE | 5914873 | 5711045 |
| MAE | 1414 | 1464 |
| MAPE | 0.2041 | 0.2296 |

Notes: *** 1%, **5%, * 10%. Robust standard errors in parentheses.

Table 3.1: Variables available from the Job Portal No.1 (aggregate data)

| | Variables | Years | Num. categories |
|---|---|---|---|
| **Job ads** | | | |
| | Occupation | 2001-2016 | 41 |
| | Industry | 2001-2016 | 41 |
| | Region | 2001-2016 | 6 |
| | Province | 2009-2016 | 77 |
| **Resumes** | | | |
| | Occupation | 2012-2016 | 41 |
| | Field of study | 2001-2016 | 44 |
| | Qualification | 2001-2016 | 3 |
| | Age | 2001-2016 | 5 |

Source: the Job Portal No.1

Table 3.2A: Variable descriptions and summary statistics of job ads from the Job Portal No.2 (micro-data)

| **Job ads (Total observations: 41,312)** | | |
|---|---|---|
| **Variables** | **Num. categories** | **% non missing** |
| Occupation | 73 | 0.00 |
| Region | 6 | 0.01 |
| Province | 76 | 0.01 |
| Salary offer (bahts) | mean of 20,315 | 47.26 |
| Work experience | 0-5 years | 56.73 |
| Qualification | Lower vocation to PhD | 90.61 |
| Num.vacancies | mean of 2.44 | 4.28 |

Source: the Job Portal No.2 (micro-data)

Table 3.2B: Variable descriptions and summary statistics of resume posts from the Job Portal No.2 (micro-data)

| **Resumes (Total observations: 89,313 )** | | |
|---|---|---|
| **Variables** | **Num. categories** | **% non missing** |
| Province (for jobs) | 76 | 0.043 |
| Gender | male, female | 0.008 |
| Age | mean of 28.5 | 0.000 |
| Reported work experience (yrs) | mean of 4.17 | 0.000 |
| Chosen start date | 4 | 0.303 |
| Requested salary (bahts) | mean of 16,500 | 99.366 |
| Occupation | 68 | 0.645 |
| Qualification | Mathayom 3 to PhD | 10.166 |
| Work status | FT, PT, freelance, intern | 0.000 |
| Grade point average | mean of 2.74 (out of 4) | 15.222 |
| Last salary | mean of 15110 | 21.868 |
| Province (of education) | 76 | 58.539 |
| Type of edu institution | 10 | 12.189 |
| Language ability | score 1-3 | 18.791 |

Source: the Job Portal No.2 (micro-data)

Table 3.3: Share of job ads with and without targeting characteristics of job candidates

| **Variables** | **Num. categories** | **% non missing** |
|---|---|---|
| Gender target | male, female, both | 0.00 |
| Age target | 15 years minimum | 74.72 |

Source: the Job Portal No.2 (micro-data)

Table 3.4: Comparing the characteristics of the Job Portal's resume data and the ICT survey

| | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-49 | 50- 59 | 60 + | Total |
|---|---|---|---|---|---|---|---|---|---|
| **(I) Individual used internet-base job search ever (ICT survey)** | 107,248 | 335,594 | 272,721 | 158,085 | 99,266 | 67,523 | 17,288 | 4,483 | 1,062,208 |
| **(II) Share of total users of internet-base job search by age (I/ Total)** | 10.10 | 31.59 | 25.67 | 14.88 | 9.35 | 6.36 | 1.63 | 0.42 | 100 |
| **(III) % of working age pop (ICT survey)** | 12.19 | 10.49 | 6.54 | 3.67 | 2.14 | 0.71 | 0.23 | 0.12 | |
| **(IV) Total Job portal resume database (in 2016)** | 1,899 | 25,479 | 28,436 | 19,266 | 9,419 | 4307 | 486 | 16 | 89,308 |
| **(V) Share of resume database by age (IV/Total)** | 2.126 | 28.529 | 31.840 | 21.573 | 10.547 | 4.823 | 0.544 | 0.018 | 100 |
| **(VI) % of resume to total ever users (IV / I)** | 1.771 | 7.592 | 10.427 | 12.187 | 9.489 | 6.379 | 2.811 | 0.357 | 8.408 |

Sources: The 2015 ICT survey (NSO), LFS 2016 Quarter 1 (NSO), job portal (accessed June/July 2016- all resumes with last update from Jan 2016-July 2016)

Table 3.5: Occupational distribution of the Job Portal's job ads data and the LD survey

| ISCO-08 1 Digit | Occupations | Vacancy Survey | Job portal No.2 micro-data | Job portal No 1 | |
|---|---|---|---|---|---|
| | | 2013 | 2016 | 2013 | 2016 |
| 1 | Managers, senior officials and legislators | 0.59 | 3.90 | 1.24 | 1.03 |
| 2 | Professionals | 7.68 | 59.20 | 38.05 | 39.36 |
| 3 | Technicians and associate professionals | 7.80 | 4.57 | 12.10 | 10.08 |
| 4 | Clerical support workers | 6.82 | 17.35 | 17.49 | 17.71 |
| 5 | Service and sales workers | 18.58 | 10.03 | 26.01 | 26.89 |
| 6 | Skilled agricultural, forestry and fishery workers | 0.26 | 0.00 | 0.00 | 0.00 |
| 7 | Craft and related trades workers | 23.53 | 4.73 | 0.14 | 0.13 |
| 8 | Plant and machine operators, and assemblers | 20.25 | 0.00 | 2.79 | 2.87 |
| 9 | Elementary occupations | 14.49 | 0.22 | 2.17 | 1.94 |

Sources: The 2013 Labour Demand Survey (NSO), micro-data Job Portal No.2, aggregated Job Portal No.1 archives

Table 3.6: Regional distribution of the Job Portal's job ads data and the LD survey

| | LD Survey | Job portal No.2 | Job portal No. 1 | |
|---|---|---|---|---|
| | 2013 | 2016 | 2013 | 2016 |
| Bangkok | 54.89 | 85.38 | 69.70 | 74.69 |
| Northern | 9.78 | 1.24 | 1.11 | 1.17 |
| Southern | 9.08 | 1.95 | 1.70 | 1.90 |
| Northeastern | 10.33 | 1.86 | 2.15 | 2.69 |
| Central | 15.91 | 9.57 | 25.34 | 19.54 |

Sources: The 2013 Labour Demand Survey (NSO), micro-data Job Portal No.2, aggregated Job Portal No.1 archives

Table 3.7: Educational distribution of the Job Portal's job ads data and the LD survey

| | **LD Survey 2013** | **Job portal No.2 2016** |
|---|---|---|
| Lower secondary or lower | 32.83 | 16.64 |
| Upper secondary, Vocational | 18.84 | 26.80 |
| Post-secondary, Higher vocational | 8.58 | 13.63 |
| Bachelor degree or higher | 15.02 | 35.86 |
| Not specify | 24.73 | 7.08 |

Sources: The 2013 Labour Demand Survey (NSO), micro-data Job Portal No.2, aggregated Job Portal No.1 archives

Table 3.8: Probability of staying at the same province

| **Dependent variables:** | **Same province of work and resident** | | | **Same region of work and residence** | | |
|---|---|---|---|---|---|---|
| Female | 0.023*** | 0.022*** | 0.024*** | 0.022*** | 0.022*** | 0.024*** |
| | [0.003] | [0.003] | [0.004] | [0.003] | [0.003] | [0.004] |
| Age 26-35 | -0.044*** | -0.047*** | -0.052*** | -0.053*** | -0.052*** | -0.055*** |
| | [0.004] | [0.005] | [0.006] | [0.004] | [0.005] | [0.006] |
| Age 36-45 | 0.001 | -0.024*** | -0.030*** | -0.025*** | -0.040*** | -0.044*** |
| | [0.005] | [0.007] | [0.008] | [0.005] | [0.007] | [0.008] |
| Age 46-55 | 0.067*** | 0.018 | 0.015 | 0.035*** | 0.004 | 0.001 |
| | [0.011] | [0.013] | [0.014] | [0.012] | [0.014] | [0.015] |
| Age 56 and above | 0.095** | 0.003 | -0.014 | 0.048 | -0.011 | -0.022 |
| | [0.043] | [0.047] | [0.050] | [0.045] | [0.048] | [0.051] |
| Currently working | | 0 | 0.001 | | 0.004 | 0.004 |
| | | [0.004] | [0.004] | | [0.004] | [0.005] |
| Experience 1-4 yrs | | 0.015*** | 0.015*** | | 0.009* | 0.009* |
| | | [0.005] | [0.005] | | [0.004] | [0.005] |
| Experience 5-10 yrs | | 0.022*** | 0.024*** | | 0.008 | 0.007 |
| | | [0.005] | [0.006] | | [0.005] | [0.006] |
| Experience 11-20 yrs | | 0.053*** | 0.052*** | | 0.037*** | 0.034*** |
| | | [0.008] | [0.008] | | [0.008] | [0.008] |
| Experience 21 yrs and more | | 0.069*** | 0.065*** | | 0.038** | 0.033* |
| | | [0.018] | [0.019] | | [0.018] | [0.019] |
| Want freelance job | | 0.072*** | 0.077*** | | 0.057*** | 0.067*** |
| | | [0.020] | [0.021] | | [0.020] | [0.021] |
| Want intern job | | 0.053 | 0.066* | | 0.060** | 0.079** |
| | | [0.034] | [0.037] | | [0.030] | [0.034] |
| Want part-time job | | 0.052*** | 0.054*** | | 0.041*** | 0.044*** |
| | | [0.012] | [0.013] | | [0.011] | [0.013] |
| Currently studying | | 0.080*** | 0.130*** | | 0.042 | 0.092** |
| | | [0.031] | [0.042] | | [0.030] | [0.042] |
| GPA | | -0.015*** | -0.014*** | | -0.014*** | -0.013*** |
| | | [0.004] | [0.004] | | [0.004] | [0.004] |
| Technical vocational | | -0.006 | -0.006 | | -0.013 | -0.011 |
| | | [0.010] | [0.011] | | [0.010] | [0.011] |
| Commerical vocational | | 0.137*** | 0.140*** | | 0.102*** | 0.109*** |
| | | [0.009] | [0.010] | | [0.009] | [0.010] |
| Rajamokol bachelor | | 0.022*** | 0.025*** | | 0.016*** | 0.018*** |
| | | [0.004] | [0.005] | | [0.005] | [0.005] |
| Rajapat bachelor | | 0.020*** | 0.023*** | | 0.035*** | 0.039*** |
| | | [0.004] | [0.005] | | [0.004] | [0.005] |
| General edu school | | 0.114*** | 0.112*** | | 0.093*** | 0.092*** |
| | | [0.012] | [0.013] | | [0.012] | [0.013] |
| Adult learning | | 0.179*** | 0.183*** | | 0.166*** | 0.173*** |
| | | [0.020] | [0.023] | | [0.018] | [0.022] |
| Skill level in English | | | 0.006* | | | 0.001 |
| | | | [0.003] | | | [0.003] |
| Observations | 94,949 | 89,895 | 77,475 | 90,553 | 85,415 | 72,697 |
| Adjusted R-squared | 0.072 | 0.086 | 0.087 | 0.124 | 0.135 | 0.123 |

Notes: ***1%, ** 5%, * 10% clustered standard errors. Other control variables are education qualification and province of work fixed effect. The data is taken from micro-data resume posting, accessed in July 2016. Omitted categories are those aged under 15, bachelor degree and no work experience.

Table 3.9 Average salary by occupation from the job-seeker's reservation level and the firm's offered level

| **Occupation** | **Reservation wage** | | **Wage offer** | |
|---|---|---|---|---|
| | **Mean** | **(SD)** | **Mean** | **(SD)** |
| Executive | 23502.13 | (18260.61) | 33799.93 | (19084.39) |
| Professional | 21240.78 | (11208.51) | 22487.45 | (12935.98) |
| Technical | 19075.07 | (8160.46) | 17864.9 | (8731.87) |
| Clerical | 16321.66 | (7038.95) | 16985.58 | (8763.45) |
| Service and sales | 16503.21 | (8368.79) | 13942.17 | (5244.74) |
| Craft and trade | 18000.42 | (9476.92) | 16038.65 | (8272.37) |
| Elementary | 15568.18 | (4048.48) | 17681.82 | (12693.08) |
| Average | 19525.79 | (10877.72) | 20315.56 | (12332.64) |

Source: micro-data Job Portal No.2

Table 3.10A: Regression for wage offer by firms

| Dependent variable: Log of wage offer | | |
|---|---|---|
| **Female only job** | -0.046*** | -0.047*** |
| | [0.007] | [0.006] |
| **Male only** | -0.092*** | -0.002 |
| | [0.008] | [0.006] |
| **Max age 26-30** | 0.082*** | 0.078*** |
| | [0.017] | [0.012] |
| **Max age over 31** | 0.152*** | 0.163*** |
| | [0.017] | [0.013] |
| **With max age target** | -0.110*** | -0.038*** |
| | [0.007] | [0.005] |
| **With experience target** | | -0.162*** |
| | | [0.007] |
| **Years of experience** | | 0.150*** |
| | | [0.004] |
| **Num. vacancies** | | -0.007*** |
| | | [0.001] |
| **Num. views** | | -0.000** |
| | | [0.000] |
| **With edu target** | | -0.032** |
| | | [0.012] |
| **LVC** | | -0.363*** |
| | | [0.008] |
| **M3** | | -0.471*** |
| | | [0.008] |
| **HVC** | | -0.268*** |
| | | [0.007] |
| **M6** | | -0.447*** |
| | | [0.008] |
| **MAS** | | 0.212*** |
| | | [0.055] |
| **Managers occ** | | 0.225*** |
| | | [0.017] |
| **Technical occ** | | -0.083*** |
| | | [0.011] |
| **Clerical Support** | | -0.133*** |
| | | [0.006] |
| **Services and Sales** | | -0.139*** |
| | | [0.007] |
| **Craft and Related Trades** | | -0.094*** |
| | | [0.011] |
| **Elementary** | | -0.156*** |
| | | [0.053] |
| **Other occ** | | -0.220*** |
| | | [0.014] |
| **Observations** | 21,311 | 19,689 |
| **Adjusted R-squared** | 0.055 | 0.507 |

Notes: ***1%, ** 5%, * 10% clustered standard errors. Other control variables are occupations and province of work fixed effect. The data is taken from micro-data resume posting, accessed in July 2016. Omitted categories are jobs omitted categories are gender-nonspecific jobs, professional occupation, bachelor degree and jobs require applicants' age under 25 yrs. All coefficients are of percentage point unit.

Table 3.10: (B) Regression for wage reservations stated by job seekers

| Dependent variable: Log of reservation wage | | |
|---|---|---|
| **Log of previous salary** | | 0.237*** |
| | | [0.003] |
| **Female** | -0.113*** | -0.073*** |
| | [0.002] | [0.002] |
| **Age 20-25** | -0.02 | -0.021 |
| | [0.014] | [0.017] |
| **Age 26-30** | -0.114*** | -0.086*** |
| | [0.014] | [0.017] |
| **Age 35 up** | -0.135*** | -0.108*** |
| | [0.014] | [0.017] |
| **On-the-job search** | 0.235*** | 0.138*** |
| | [0.002] | [0.003] |
| **Exp yrs: 1-4** | 0.118*** | -0.052*** |
| | [0.002] | [0.004] |
| **Exp yrs: 5-10** | 0.398*** | 0.078*** |
| | [0.003] | [0.005] |
| **Exp yrs: 11-20** | 0.657*** | 0.237*** |
| | [0.005] | [0.007] |
| **Exp yrs: 20 above** | 0.817*** | 0.330*** |
| | [0.018] | [0.016] |
| **Freelance** | -0.036** | -0.016 |
| | [0.017] | [0.021] |
| **Intern** | -0.353*** | -0.198*** |
| | [0.039] | [0.055] |
| **Part-time** | -0.327*** | -0.247*** |
| | [0.009] | [0.013] |
| **At education** | 0.081*** | 0.032 |
| | [0.022] | [0.026] |
| **Ed: M3 current** | -0.369*** | -0.289*** |
| | [0.008] | [0.008] |
| **Ed: M3** | -0.413*** | -0.337*** |
| | [0.008] | [0.008] |
| **Ed: HVC current** | -0.321*** | -0.253*** |
| | [0.030] | [0.037] |
| **Ed: HVC** | -0.278*** | -0.219*** |
| | [0.006] | [0.006] |
| **Ed: M6** | -0.330*** | -0.275*** |
| | [0.007] | [0.007] |
| **Ed: Bach current** | -0.236*** | -0.189*** |
| | [0.023] | [0.028] |
| **Ed: Master current** | 0.162*** | 0.141*** |
| | [0.027] | [0.031] |
| **Ed: Master** | 0.324*** | 0.217*** |
| | [0.006] | [0.006] |
| **Ed: PhD current** | 0.834*** | 0.730*** |
| | [0.171] | [0.185] |
| **Ed: Phd** | 0.16 | 0.322*** |
| | [0.173] | [0.101] |

| | | |
|---|---|---|
| **GPA** | 0.027*** | 0.034*** |
| | [0.002] | [0.002] |
| **Technical vocational** | -0.042*** | -0.028*** |
| | [0.006] | [0.006] |
| **Commercial vocational** | -0.061*** | -0.046*** |
| | [0.005] | [0.005] |
| **Rajamongkol bachelor** | 0.055*** | 0.051*** |
| | [0.003] | [0.003] |
| **Rajapat bachelor** | -0.093*** | -0.068*** |
| | [0.002] | [0.002] |
| **General edu school** | -0.118*** | -0.083*** |
| | [0.008] | [0.008] |
| **Adult learning** | -0.136*** | -0.096*** |
| | [0.010] | [0.010] |
| **Know a European language** | 0.016 | -0.002 |
| | [0.025] | [0.038] |
| **Know an Asian language** | 0.097*** | 0.094*** |
| | [0.009] | [0.013] |
| **Skill level in English** | 0.059*** | 0.038*** |
| | [0.002] | [0.002] |
| **Observations** | 152,429 | 108,089 |
| **Adjusted R-squared** | 0.485 | 0.603 |

Notes: ***1%, ** 5%, * 10% clustered standard errors. Other control variables are education qualification and province of work fixed effect. The data is taken from micro-data resume posting, accessed in July 2016. Omitted categories are those aged under 15, bachelor degree and no work experience. Column II also controls for occupation-type.

Table 3.11: Share of total job ads that target certain traits of applicants

| | Percentage |
|---|---|
| Gender targeting | |
| Only female | 14.47 |
| Only male | 17.93 |
| Any gender | 67.61 |
| Gender missing | 0 |
| Age targeting | |
| 16-19 | 9.68 |
| 20-14 | 38.39 |
| 25-29 | 19.5 |
| 30 or above | 7.03 |
| Any age | 25.4 |
| Total | 41,312 |

Source: micro-data Job Portal No.2