

# Measuring Vacancies: Firm-level Evidence from Two Measures\*

Niels-Jakob Harbo Hansen<sup>1</sup> and Hans Henrik Sievertsen<sup>†2</sup>

<sup>1</sup>Institute for International Economic Studies (IIES)

<sup>2</sup>The Danish National Centre for Social Research (SFI)

This version: August 8, 2016

## Abstract

Using firm-level survey- and register-data for both Sweden and Denmark we show systematic mis-measurement in both vacancy measures. While the register-based measure on the aggregate constitutes a quarter of the survey-based measure, the latter is not a super-set of the former. To obtain the full set of unique vacancies in these two databases, the number of survey vacancies should be multiplied by approximately 1.2. Importantly, this adjustment factor varies over time and across firm characteristics. Our findings have implications for both the search-matching literature and policy analysis based on vacancy measures: Observed changes in vacancies can be an outcome of changes in mis-measurement, and are not necessarily changes in the actual number of vacancies.

JEL codes: J23, J60

Keywords: Vacancies, Surveys, Public Employment Services

---

\*We are indebted to Orley C. Ashenfelter, Paul Bingley, Per Krusell, Karsten Albæk, John Hassler, Robert Shimer, Tobias Broer, Hannes Malmberg, Georg Marthin and Erik Öberg for stimulating discussions and suggestions. We are very grateful to Charlotte Leolnar Reif, Anders Gabriel Pedersen, Klaus Henrik Langager and Charlotte Funk Christensen for providing us with data. All remaining errors are our own. Sievertsen acknowledges financial support from the Danish Council for Independent Research through grant 09-070295. Hansen acknowledges financial support from Institute for Evaluation of Labour Market and Education Policy and Handelsbanken's Research Foundations.

<sup>†</sup>Corresponding author. Address: SFI, Herluf Trolles Gade 11, 1052, Copenhagen. Denmark. Email: [hhs@sfi.dk](mailto:hhs@sfi.dk).

# 1 Introduction

In the empirical search-matching literature job openings are often measured using two alternative sources. The first is survey data on job vacancies compiled by statistical agencies. The second is register data on vacancies from databases maintained by Public Employment Services (PES). The former is often considered preferable, as it avoids the selection problems that are likely to be present in the latter (Elsby et al., 2015). However, since survey data for many countries, particularly in Europe, so far only is available for relatively short time periods the data from public job centers are used as proxy in many studies (Berman, 1997; Carlsson et al., 2013; Albæk and Hansen, 2004; Wall and Zoega, 2002; Yashiv, 2000).

An important, yet unanswered, question is how these two measures relate to each other on the micro level. This is important for a number of reasons. First, if we are interested in using the register based measure as proxy for the survey on a subset of the labor market, for which the survey is not available, we cannot simply rely on how well the aggregate times series track each other. Second, it is often assumed that the register based measure suffers from selection problems Elsby et al. (2015) write “a perennial concern is that such registered vacancies fail to be representative of all job openings”. These problems are likely to be smaller in the survey based measure. Yet it is also possible that the survey measure is not a super-set of the vacancies in the PES. That is, some vacancies are registered at the PES but not in the survey. Thus, by comparing the joint distribution on the micro-level of the two vacancy measures, we can obtain important information about the reliability of the data sources.

We investigate this question using firm level data for Sweden and Denmark. Specifically, we construct two new databases on the firm level. In the first database we match data for vacancies as measured in a survey conducted by Statistics Sweden with data for vacancies as counted in the database by the Swedish Public Employment Service. This database covers the period 2001-2012. In the second database we match data for vacancies as measured in a survey conducted by Statistics Denmark, with data for vacancies as counted in the database by the Danish Public Employment Service. This database covers

the period 2010-2012. Using these databases we relate the two measures to each other across time and firm level characteristics.

We find that on the the aggregate level, the number of vacancies in the PES accounts for roughly a quarter of all vacancies in the survey. This does not, however, imply that the vacancies in the survey constitute a superset of the PES vacancies. Specifically, we find that about 60% of the vacancies in the PES are not included in the survey. To obtain the number of unique vacancies in the survey and PES, the survey needs to be adjusted by a factor of around 1.2. Furthermore, we find that this correction factor varies the over business cycle and with firm characteristics. This is important for both research on business cycles as it means that fluctuations in survey vacancies could be driven by changes in mis-measurement.

These findings are also important in a methodological perspective. There appears to be consensus that surveys represent the best available measures of vacancies (Elsby et al., 2015; Shimer, 2005).<sup>1</sup> However, our results suggest that reliance on survey measure is problematic. The extra 20% of unique vacancies present in the PES provides a lower bound vacancies missing from the survey measure. This is one potential reason why a non-trivial share of hires are made with survey vacancies presented in the previous month. This phenomenon is documented by Davis et al. (2013) in the United States, and by Chapter 2 of this thesis for Sweden.

The paper proceeds as follows. In Section 2, we describe our data sources and how the Danish and Swedish database are constructed. Section 3 documents the relationship between the two vacancy measures on the firm level. In Section 4 we compute the aggregate number of unique vacancies in the two databases and relate these to the aggregate number of vacancies in the survey. Section 5 concludes.

---

<sup>1</sup>Shimer (2005) argues that the vacancy survey in the U.S. (JOLTS) provides “an ideal empirical definition” of job vacancies. Elsby et al. (2015) writes “Reliable, timely and comprehensive survey data have only recently been made available since the inception of the Job Openings and Labor Turnover Survey (JOLTS).”

## 2 Data

In this section we describe our data sources and how we construct our two new databases. We first describe our data on vacancies from surveys and then describe our data on vacancies measured by the Public Employment Services. Below we will refer to the first data source as *survey data* and the second as *register data*.

### 2.1 Data from vacancy surveys

Our first data sources for vacancies are the Danish and Swedish *Job Vacancy Survey*, respectively.

#### Sweden

The *Swedish Job Vacancy Survey* is administered by Statistics Sweden and has been collected on a quarterly basis since 2001. Two vacancy concepts are available in this survey: (1) the number of available positions in each plant, which have been made “available for external job-seekers via the newspapers, internet or another mean of dissemination”, and (2) the number of these positions that the employer wishes to fill immediately.<sup>2</sup> This way, the former concept is a super-set of the latter. Below we rely on (1).

The *Swedish Job Vacancy Survey* is collected at the plant level and all respondents are asked to report the number of vacancies in the middle of the reference month.<sup>3</sup> For the private sector the sampling is carried out on the plant level with approximately 16,700 work places sampled each period. For the public sector the sampling was also conducted on the plant level until the second quarter of 2006, when the sampling was changed to the organizational level and on this level 650 organizations are sampled each period. Units larger than 100 employees are asked to do the reporting for each month of the relevant quarter, whereas units with less than 100 employees only are asked to report in the reference month. Reporting takes place either via letter or online. Non-respondents are reminded via email, letter, or a phone call. Until 2004, reporting was voluntary and

---

<sup>2</sup>In Swedish (1) is called *Vakanser* and (2) is called *Lediga jobb*.

<sup>3</sup>The respondents are asked to report the number of job openings on the Wednesday closest to the 15th of the reference month.

the share of non-reporting units roughly 30%. In 2004 reporting became mandatory, and currently the share of non-reporting units is 11% in the private sector and 2% in the public sector.

## Denmark

The *Danish Job Vacancy Survey* is administered by Statistics Denmark, and has been collected on quarterly basis since 2010 (Danmarks Statistik, 2016). We have access to the data for the period 2010-12. The vacancy concept in this survey is a “position with salary, which is newly established, idle or about to be vacant, and that the employer is taking active steps to fill with a suitable candidate from outside the firm”. The plant is asked to report its number of vacancies on the second Wednesday of the given month.

The sampling is conducted on the plant level, with 7,000 plants surveyed each quarter. The plants are sampled from the *Central Business Register*, which is a database covering all registered firms in Denmark. Data is reported on a quarterly basis but collected monthly with three equally sized groups reporting each month in the relevant quarter. Reporting firms are selected into one of these three groups randomly. Reporting is mandatory and plants that fail to report are first reminded and then informed that lack of reporting will lead to a police report. The share of non-reporting units is 5%. For non-reporting units larger than 100 employees data is imputed.<sup>4</sup>

That data is reported quarterly, but collected monthly, which constitutes a challenge. Indeed, for a given observation we know only the quarter for which the data was collected, while the exact month is unknown to us. We thus assume that the observed number of observations is representative for all months in the particular quarter. Specifically, we compare the observed number of survey vacancies to the *average* number of PES vacancies in the relevant quarter.

---

<sup>4</sup>Imputation is done for 2.5% of all plants. This is accomplished via prediction from a regression model of vacancies on the number of employees within the same industry.

## 2.2 Data from Public Employment Service

Our second data source for vacancies is the Danish and Swedish register data from the Public Employment Service (PES).

### Sweden

For Sweden, we have access to all vacancies registered at the *Swedish Public Employment Service*. This database covers the all job vacancies made at the agency during the period 2001-2012. The database contains a row for each posting made at the agency with information on the start and end date of the posting along with information on the number of workers the firm is searching for and information on job and firm type. In principle, the database contains both a firm and plant identifier. However, for the majority of the observations the plant identifiers are missing, so the database in practice only contains useful information on the firm level. The database is known to contain a number of duplicates. To remove these, we drop vacancies that are identical to another vacancy in firm identifier, start-date, and job-characteristics.<sup>5</sup> Furthermore, to allow for comparison to the survey, we restrict our sample to the PES vacancies that are active on the reference day of the survey.

### Denmark

For Denmark, we use all vacancies announced at the *Danish Public Employment Service* throughout the period 2010-2012. Each row in the database contains a firm identifier, the number of positions available, on the announcement day of the vacancy as well as the date on the vacancy was withdrawn from the database. As in the Swedish case we limit our sample to vacancies active on the reference day of the survey.

## 2.3 Background variables

For both Denmark and Sweden, we have access to background variables at the firm level.

---

<sup>5</sup>This operation removes 8% of the rows in the database. One row can contain more than one vacancy.

For Sweden we have access to information on industry, firm size, value added, and turnover. The vacancy survey provides industry information on the NACE level, and using this data we classify the firms into eight industry groups.<sup>6</sup> We further have access to a subset of the variables from the company section of the *Longitudinal Integration Database for Health Insurance and Labour Market Studies* (LISA). This is a register database maintained by Statistics Sweden, that provides information on firm size, value added and turnover for the registered firms at an annual frequency.

For Denmark we have access to industry information, firm size and average wage level on the firm level. All these data derive from *The Integrated Database for Employment Research* (IDA), which is developed and maintained by Statistics Denmark. The database contains the set of firms where at least one person is registered as being employed during November each year.

## 2.4 Data selection

The survey data is collected on the plant level, while the register data only has useful identifiers on the firm level. To address this issue, we restrict our sample to firms with only one plant. For Denmark we obtain the number of plants per firm from the IDA database. For Sweden the number is obtained from the IFAU database, which contains information on all employment spells on the plant and firm level. This data is based on Swedish tax records, and allows us to compare the correct number of survey and register vacancies to each other, but it clearly comes at the cost of restricting attention to a non-random subset of the data. In Table 1, we show that 43 and 35% of all observations for Sweden and Denmark, respectively, are lost through this operation.

Analyses that exploit firm background variables further reduces the size of the sample. While the reduction is negligible in DK, the sample is reduced by 14% in the Swedish case.

In Table A1 we assess the representativeness of our sample *vis-a-vis* the distributions of firms in the wider economy. Table A1 shows that *Farming, fishery, and mining* is

---

<sup>6</sup>The Statistical classification of economic activities in the European Community is abbreviated as NACE.

over-sampled in both countries, while *manufacturing* is under-sampled. Besides that, the distribution across industries roughly mimics that in the economy. Across size, our sample has an over-representation of larger firms, while firms with high value added (for Sweden)/wage (for Denmark) per worker are over-sampled. This may seem at odds with the data restriction to firms with only one plant, but larger firms are likely to have been over-sampled initially in the vacancy surveys.

Finally, for a fraction of observations in the PES data there are missing firm identifiers. For Denmark this share is low (1.5%) over the entire sample period, while for Sweden it falls from around 30% in 2001 to 1% in 2010 and remains constant thereafter.<sup>7</sup> However, our conclusions are robust to restricting the Swedish sample to the period after 2009.<sup>8</sup>

**Table 1:** Sample selection

|                            | SWE     |      |     | DEN    |      |     |
|----------------------------|---------|------|-----|--------|------|-----|
|                            | N       | Frac | Vac | N      | Frac | Vac |
| Full sample                | 852,156 | 1.00 | 2.0 | 63,722 | 1.00 | 1.2 |
| Information on # of plants | 726,058 | 0.85 | 2.1 | 61,496 | 0.97 | 1.3 |
| One plant per firm         | 487,667 | 0.57 | 0.8 | 41,731 | 0.65 | 0.5 |
| Covariates available       | 366,610 | 0.43 | 0.9 | 41,668 | 0.65 | 0.5 |

## 2.5 Descriptive statistics

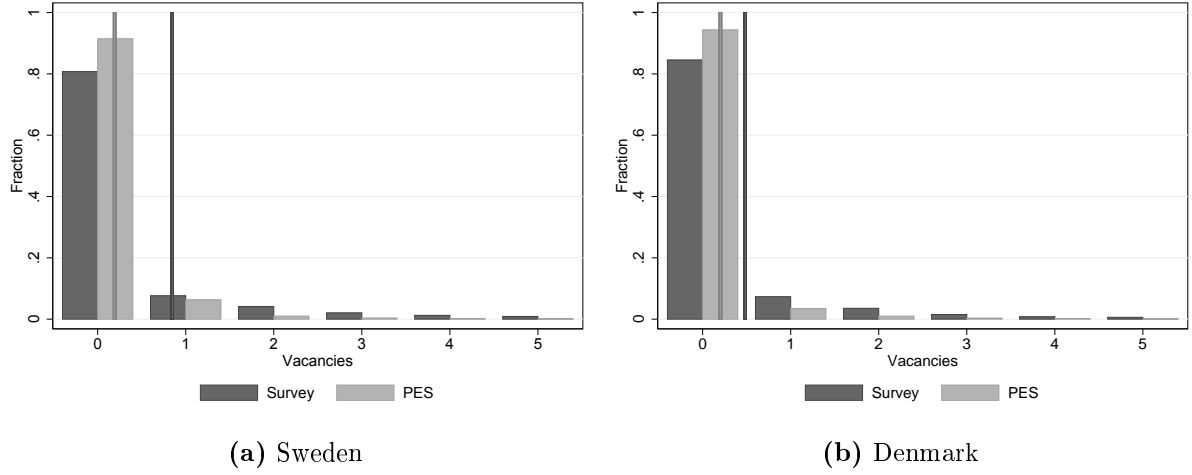
Figure 1 shows the distribution of vacancies for both measures and countries. We note three similarities between the two countries. One, is that more than 80% of the observations have zero vacancies. A second is that the fraction of observations with zero vacancies is somewhat higher in the PES than in the survey. The third similarity is that the fraction of observations is monotonically decreasing in the number of vacancies. There are, however, also a dissimilarity between the countries: the average of survey vacancies is 75% higher in the Swedish data.

Table 2 presents means and standard deviations for the background variables. The

<sup>7</sup>The exact shares are 32.6% in 2001, 26.2% in 2002, 17.8% in 2003, 13.25% in 2004, 10.5% in 2005, 8.4% in 2006, 6.6% in 2007, 3.4% in 2008, 2.8% in 2009, 1.1% in 2010, 1.2% in 2011 and 1.1% in 2012.

<sup>8</sup>To check for the impact of this issue we have redone our main table for Sweden (Table 6) with a sample restricted to after 2009 in the Table A2.





**Figure 1:** Vacancy distribution in the Survey and the Public Employment Services data (PES)

distribution of firms across both industries and sizes differs considerably across the two samples.

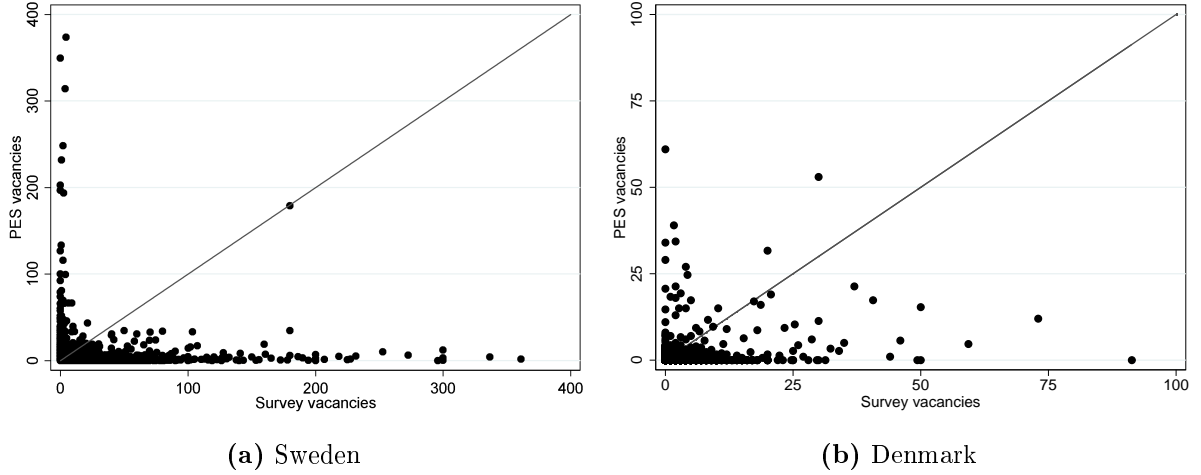
**Table 2:** Descriptive Statistics

|                                       | SWE  |       | DEN  |      |
|---------------------------------------|------|-------|------|------|
|                                       | Mean | SD    | Mean | SD   |
| <i>Industry</i>                       |      |       |      |      |
| Farming, fishery and mining           | 0.06 | 0.24  | 0.03 | 0.18 |
| Manufacturing                         | 0.16 | 0.37  | 0.15 | 0.36 |
| Energy and water supply               | 0.01 | 0.09  | 0.15 | 0.35 |
| Construction                          | 0.09 | 0.29  | 0.31 | 0.46 |
| Trade, Hotel and Restaurants          | 0.20 | 0.40  | 0.11 | 0.31 |
| Transportation, mail and telecom      | 0.05 | 0.22  | 0.05 | 0.23 |
| Finance and business services         | 0.25 | 0.43  | 0.14 | 0.35 |
| Public services and personal services | 0.17 | 0.38  | 0.06 | 0.24 |
| <i>Firm size</i>                      |      |       |      |      |
| No. of employees                      | 16.6 | 43.7  | 26.2 | 56.4 |
| 0-5                                   | 0.30 | 0.46  | 0.42 | 0.49 |
| 6-10                                  | 0.11 | 0.31  | 0.14 | 0.35 |
| 11-50                                 | 0.16 | 0.36  | 0.29 | 0.45 |
| 50+                                   | 0.43 | 0.50  | 0.14 | 0.35 |
| <i>Value-added/salary</i>             |      |       |      |      |
| Value-added per worker (1,000 EUR)    | 64.1 | 134.0 |      |      |
| Gross salary per worker (1,000 EUR)   |      |       | 44.4 | 51.3 |

Notes: Value-added per worker and gross salary per worker is adjusted to the 2012 price level using the harmonized consumer price index for both countries and converted to EUR using the 2012 exchange rates.

### 3 Patterns at the firm level

Figure 2 shows scatter plots of vacancies as measured in the survey and in the PES on the firm level. One dot represents the number of vacancies in the survey. For the PES, a dot is an average over three firms for one point in time.<sup>9</sup> Two observations are clear from this figure. First, the dots are not scattered around the 45-degree line, which would be the case if the two measures represented the same vacancy concept. Second, there are a substantial number of observations above the 45-degree line. This is surprising as one would expect the survey data to be a super-set of the data from the PES, following the definitions in Section 2.



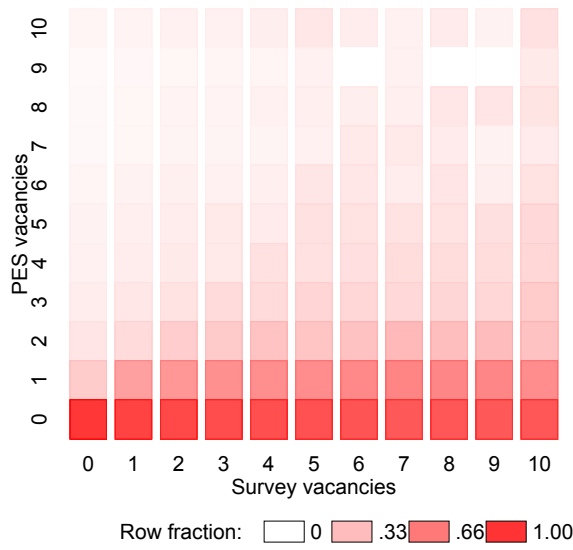
**Figure 2:** PES vacancies plotted against survey vacancies for Sweden and Denmark. Each dot represents the average across three firms.

The interpretation of the scatter plot could be clouded by the substantial number of observations around zero. However, regression results confirm that the slope coefficient is considerably smaller than one.<sup>10</sup>

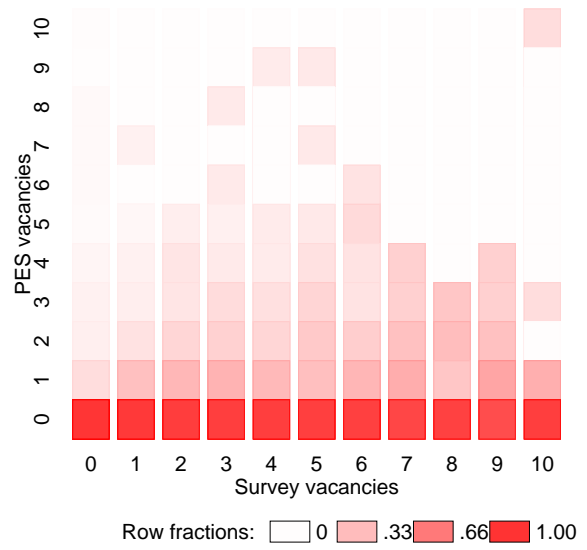
In Figure 3 we look more closely at the joint distribution of the PES and survey vacancies. Panels (a) and (b) show the concentration of PES vacancies for a given number of survey vacancies. From these graphs we see that regardless of the number of observations registered in the survey, most of the density mass is concentrated at zero PES vacancies. However, the variation in PES vacancies increases in the number of survey vacancies.

<sup>9</sup>The averaging is carried for anonymity reasons, but does not affect the patterns in the figures.

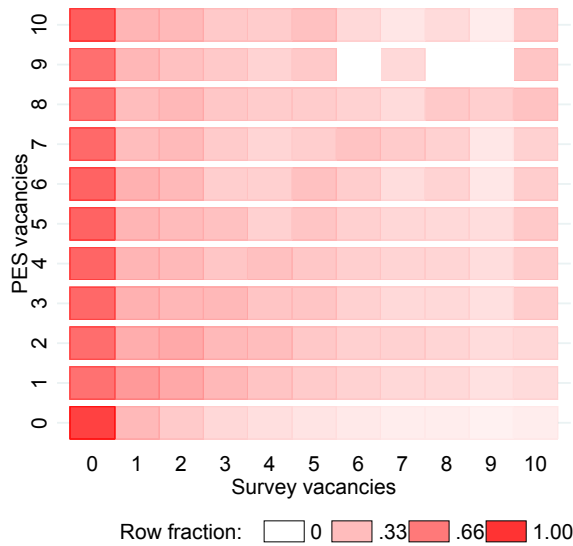
<sup>10</sup>These regressions are available upon request.



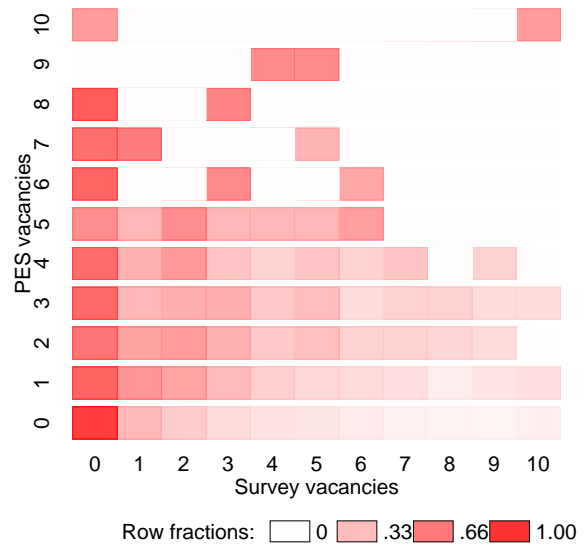
(a) By survey columns, Sweden



(b) By survey columns, Denmark



(c) By PES rows, Sweden



(d) By PES rows, Denmark

**Figure 3:** Distributions of PES vacancies for given levels of survey vacancies and vice versa.

Panels (c) and (d) show the same graphs, but with the concentration of survey vacancies for a given number of PES vacancies. A similar pattern is observed here. The pattern is clearer for Sweden, which is likely due to the larger sample size. Table 3 also presents the joint distribution. We note that around 80% of the observations have both zero PES and survey vacancies, 18% (15%) of the observations have more survey than PES vacancies in Sweden (Denmark), and 4% (2%) have more PES than survey vacancies.

**Table 3:** Row and Column Elements

|     |    | Sweden |       |       |       |       |  |    | Denmark |       |       |       |       |
|-----|----|--------|-------|-------|-------|-------|--|----|---------|-------|-------|-------|-------|
|     |    | Survey |       |       |       |       |  |    | Survey  |       |       |       |       |
|     |    | 0      | 1     | 2     | 3     | 4+    |  |    | 0       | 1     | 2     | 3     | 4+    |
| PES | 0  | 0.771  | 0.063 | 0.032 | 0.015 | 0.033 |  | 0  | 0.832   | 0.069 | 0.031 | 0.013 | 0.025 |
|     | 1  | 0.026  | 0.011 | 0.008 | 0.005 | 0.014 |  | 1  | 0.010   | 0.004 | 0.003 | 0.001 | 0.002 |
|     | 2  | 0.004  | 0.001 | 0.001 | 0.001 | 0.003 |  | 2  | 0.002   | 0.001 | 0.001 | 0.000 | 0.001 |
|     | 3  | 0.002  | 0.000 | 0.000 | 0.000 | 0.001 |  | 3  | 0.001   | 0.000 | 0.000 | 0.000 | 0.001 |
|     | 4+ | 0.004  | 0.001 | 0.001 | 0.000 | 0.002 |  | 4+ | 0.000   | 0.000 | 0.000 | 0.000 | 0.000 |

In Table 4 we show the ratio of PES to survey vacancies (column  $P/S$ ). Columns  $PS$  and  $SS$  in the table show the relevant shares by subgroups of total PES and survey vacancies, respectively. Four observations can be made from this table. First, the aggregate share of PES to survey vacancies is around one quarter in both countries. Second, while *Farming, fishery and mining* is the industry with the highest fraction in Sweden, *Public and personal services* is the industry with the highest fraction in Denmark. Note however, that *Farming, fishery and mining* only accounts for a small share of both survey and PES vacancies. Third, for Sweden the fraction of PES to survey vacancies is decreasing in firm size, while no clear pattern is visible in the Danish data. Fourth, for Sweden we observe that the share is decreasing in value added per worker, while for Denmark the share is decreasing in salary per worker.

Table 4 does not highlight that a given aggregate fraction of PES to survey vacancies does not need to stem from the same fraction in all firms. Thus, in Table 5 we look more closely at the relationship between survey and PES vacancies at the firm level, as first presented in Figure 2. In columns (1) and (5) we report the number of observations where the two measures are identical. Columns (2) and (6) are similar but involve a restriction the sample to observations where at least one of two measure is non-zero. In columns (3) and (7) we report the fraction of observations with a higher number of PES than survey vacancies. Columns (4) and (8) report the fraction of all PES vacancies that do not have corresponding number of survey vacancy registered. The fractions are computed in relation to all survey vacancies: *E.g.* if a firm has posted 10 PES and 2 survey vacancies in one period, then this firm contribute with 8 to the numerator of the NC measure in

**Table 4:** Fractions

|   | SWE        |           |           | DEN        |           |           |
|---|------------|-----------|-----------|------------|-----------|-----------|
|   | <i>P/S</i> | <i>PS</i> | <i>SS</i> | <i>P/S</i> | <i>PS</i> | <i>SS</i> |
| All                                     | 0.23       | 1.00      | 1.00      | 0.23       | 1.00      | 1.00      |
| <i>By industry</i>                      |            |           |           |            |           |           |
| Farming, fishery and mining             | 0.63       | 0.02      | 0.01      | 0.09       | 0.01      | 0.03      |
| Manufacturing                           | 0.20       | 0.30      | 0.33      | 0.10       | 0.11      | 0.25      |
| Energy and water supply                 | 0.10       | 0.01      | 0.02      | 0.09       | 0.02      | 0.06      |
| Construction                            | 0.21       | 0.02      | 0.02      | 0.08       | 0.06      | 0.17      |
| Trade, Hotel and Restaurants            | 0.25       | 0.11      | 0.10      | 0.36       | 0.10      | 0.07      |
| Transportation, mail and telecom        | 0.21       | 0.06      | 0.06      | 0.06       | 0.03      | 0.12      |
| Finance and business services           | 0.23       | 0.23      | 0.23      | 0.33       | 0.19      | 0.13      |
| Public services and personal services   | 0.24       | 0.25      | 0.23      | 0.74       | 0.46      | 0.14      |
| <i>By firm size</i>                     |            |           |           |            |           |           |
| 0-5                                     | 0.38       | 0.02      | 0.01      | 0.13       | 0.03      | 0.05      |
| 6-10                                    | 0.38       | 0.03      | 0.02      | 0.21       | 0.03      | 0.03      |
| 11-50                                   | 0.37       | 0.13      | 0.08      | 0.45       | 0.34      | 0.18      |
| 50+                                     | 0.21       | 0.82      | 0.89      | 0.19       | 0.59      | 0.74      |
| <i>By value-added/salary per worker</i> |            |           |           |            |           |           |
| Quartile 1                              | 0.35       | 0.23      | 0.15      | 0.82       | 0.50      | 0.14      |
| Quartile 2                              | 0.36       | 0.20      | 0.13      | 0.46       | 0.24      | 0.12      |
| Quartile 3                              | 0.20       | 0.17      | 0.19      | 0.13       | 0.12      | 0.22      |
| Quartile 4                              | 0.17       | 0.22      | 0.29      | 0.05       | 0.10      | 0.49      |

Notes: *P* : PES vacancies, *S* : survey vacancies, *PS* : share of all PES vacancies, and *SS* : share of all survey vacancies. For Sweden the quartiles refer to value-added per worker. For Denmark the quartiles refer to the gross salary per worker.

columns (4) and (8).

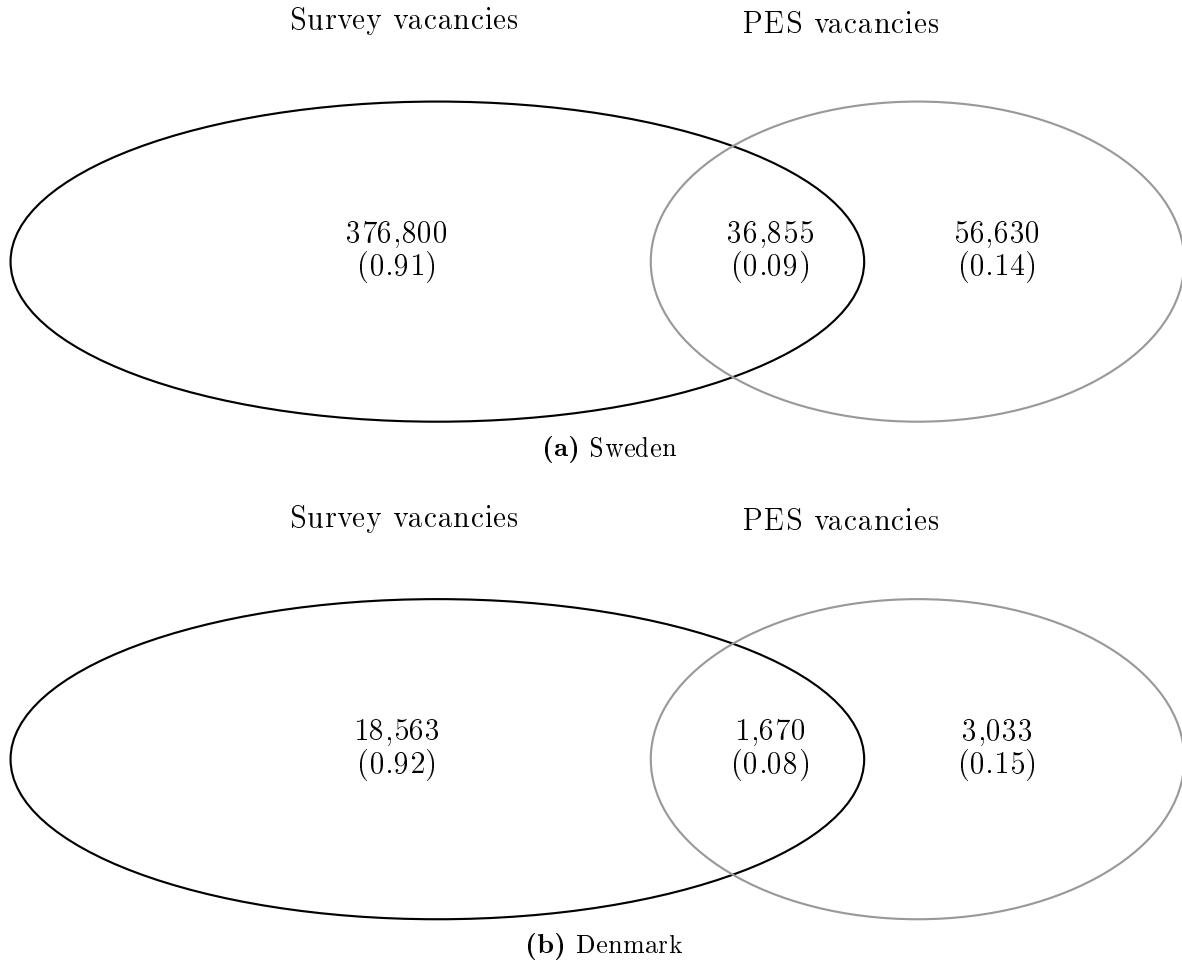
In both countries the share of observations with an equal number of PES and survey vacancies is around 80%. However, a large share of these observations are located at the origin. Indeed, only 7% and 3% of all non-zero observations are identical. As noted in the discussion of Figure 2 above, there are also observations with a higher number of PES than survey vacancies. These account for 4% of all observations in Sweden and 2% in Denmark.

The findings so far have shown that the aggregate number of PES vacancies constitutes a fraction of the survey vacancies. This may lead one to presume that PES vacancies simply form a subset of survey vacancies. In column (4) and (8) we show that this presumption is false. Indeed, 61% and 64% of the vacancies in the PES are not present in the survey.

**Table 5:** Fractions on line

|   | SWE            |                |                |             | DEN            |                |                |             |
|---|----------------|----------------|----------------|-------------|----------------|----------------|----------------|-------------|
|   | (1)<br>$P = S$ | (2)<br>$P = S$ | (3)<br>$P > S$ | (4)<br>$NC$ | (5)<br>$P = S$ | (6)<br>$P = S$ | (7)<br>$P > S$ | (8)<br>$NC$ |
| All                                       | 0.78           | 0.07           | 0.04           | 0.61        | 0.84           | 0.03           | 0.02           | 0.64        |
| <i>By industry</i>                        |                |                |                |             |                |                |                |             |
| Farming, fishery and mining               | 0.93           | 0.06           | 0.01           | 0.83        | 0.84           | 0.03           | 0.02           | 0.76        |
| Manufacturing                             | 0.72           | 0.07           | 0.04           | 0.57        | 0.76           | 0.03           | 0.01           | 0.40        |
| Energy and water supply                   | 0.75           | 0.07           | 0.04           | 0.49        | 0.91           | 0.02           | 0.01           | 0.61        |
| Construction                              | 0.90           | 0.04           | 0.01           | 0.52        | 0.87           | 0.02           | 0.01           | 0.59        |
| Trade, Hotel and Restaurants              | 0.82           | 0.08           | 0.04           | 0.65        | 0.84           | 0.04           | 0.03           | 0.82        |
| Transportation, mail and telecom          | 0.77           | 0.04           | 0.04           | 0.65        | 0.68           | 0.01           | 0.01           | 0.54        |
| Finance and business services             | 0.79           | 0.05           | 0.04           | 0.63        | 0.81           | 0.03           | 0.01           | 0.86        |
| Public services and personal services     | 0.76           | 0.10           | 0.06           | 0.59        | 0.77           | 0.04           | 0.09           | 0.60        |
| <i>By firm size</i>                       |                |                |                |             |                |                |                |             |
| 0-5                                       | 0.97           | 0.06           | 0.01           | 0.82        | 0.96           | 0.01           | 0.00           | 0.67        |
| 6-10                                      | 0.91           | 0.08           | 0.02           | 0.74        | 0.92           | 0.02           | 0.01           | 0.79        |
| 11-50                                     | 0.83           | 0.08           | 0.04           | 0.73        | 0.84           | 0.03           | 0.02           | 0.82        |
| 50+                                       | 0.61           | 0.06           | 0.07           | 0.58        | 0.56           | 0.03           | 0.04           | 0.54        |
| <i>By value-added per worker quartile</i> |                |                |                |             |                |                |                |             |
| Quartile 1                                | 0.83           | 0.08           | 0.06           | 0.64        | 0.91           | 0.04           | 0.03           | 0.68        |
| Quartile 2                                | 0.81           | 0.09           | 0.05           | 0.65        | 0.90           | 0.05           | 0.02           | 0.71        |
| Quartile 3                                | 0.77           | 0.07           | 0.04           | 0.57        | 0.81           | 0.03           | 0.01           | 0.46        |
| Quartile 4                                | 0.72           | 0.05           | 0.03           | 0.61        | 0.71           | 0.01           | 0.01           | 0.52        |
| Vacancy sample                            | All            | $S > 0$        | All            | All         | All            | $S > 0$        | All            | All         |

Notes:  $P$  : PES vacancies,  $S$  : survey vacancies and  $NC$  : Fraction of PES vacancies with no survey vacancy. (1) and (5) report the number of observations where the two measures are identical, columns (2) and (6) do the same but with the sampled restricted to observations where at least one of two measure is non-zero. Columns (3) and (7) report the fraction of observations with a higher number of PES than survey vacancies. Columns (4) and (8) reports the fraction of all PES vacancies, which do not have a number corresponding survey vacancy registered as a share of all survey vacancies.



**Figure 4:** Vacancies in the Survey, the PES and total vacancies. The uppermost numbers are the number of vacancies. The numbers in parenthesis indicate the fraction relative to the total number of survey vacancies.

Figure 4 shows the overlap and dis joint sub-sets of the PES and survey vacancies. Slightly more than 90% of the survey vacancies are not present in the PES database and the remaining 10% are present in both databases. Finally, the number of survey vacancies would grow by approximately 15% if all PES vacancies were also present in the survey. That is, if the survey truly was a super-set of the PES vacancies.

To assess whether the ratio of total vacancies to survey vacancies varies systematically across firms we regress the ratio on observable firm characteristics in Table 6. Columns (2) and (4) show that for both countries the ratio increases in firm size. Also for both countries firms within *Public and personal services* have significantly higher ratios than for the reference group, *Energy and water supply*.

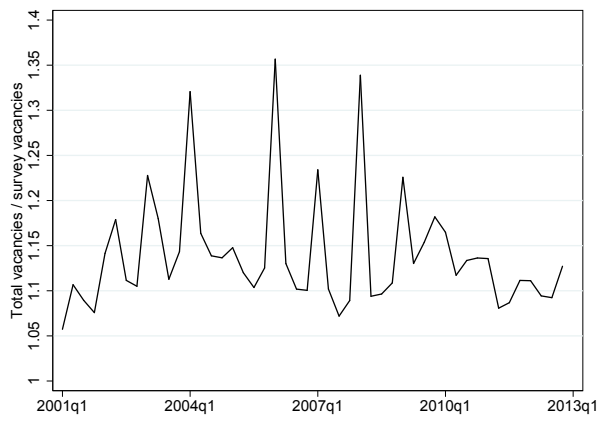
## 4 Aggregate implications

Above we have shown that a substantial amount of vacancies are present in the PES but not in the survey data. This means, that one cannot simply interpret the latter as a super-set of the former. This has implications on the aggregate level, as one has to take into account the vacancies present in the PES, but not in the survey, in order to get closer to the true number of vacancies in the economy.

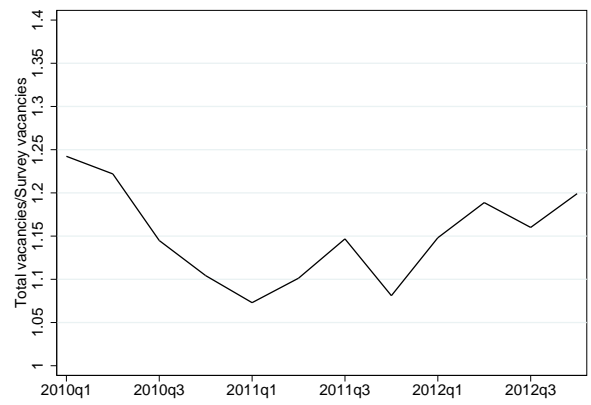
In Figure 5 we report the full set of unique vacancies in our database to the number of vacancies counted in the survey. The full set of unique vacancies is computed by adding the vacancies unique to the PES to the total number of survey vacancies. For both countries, this ratio varies in the range 1 to 1.3. In Appendix Figure A1 we report the underlying times series.

To assess whether the ratio varies systematically over the business cycle, Table 6 reports the coefficients from a regression of the ratio on the rate of unemployment. For Sweden the ratio varies negatively with the rate of unemployment, while this relationship is insignificant in Denmark. The latter is likely due to a shorter time period. In other words, when labor market conditions improve the level of mis-measurement tends to increase.





(a) Sweden



(b) Denmark

**Figure 5:** Aggregated total vacancies related to aggregated survey vacancies

**Table 6:** Regressions results - Dependent variable: The ratio of total vacancies to survey vacancies

|                                       | SWE               |                   |                   | DEN               |                   |                   |
|---------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                                       | (1)               | (2)               | (3)               | (4)               | (5)               | (6)               |
| Unemployment                          | -0.14<br>(0.18)   |                   | -0.82**<br>(0.36) | -0.01<br>(0.48)   |                   | 0.02<br>(0.50)    |
| Farming, fishery and mining           |                   | 0.02**<br>(0.01)  | 0.02**<br>(0.01)  |                   | -0.01<br>(0.01)   | -0.01<br>(0.01)   |
| Manufacturing                         |                   | 0.02<br>(0.02)    | 0.02<br>(0.02)    |                   | -0.00<br>(0.00)   | -0.00<br>(0.00)   |
| Construction                          |                   | 0.01<br>(0.01)    | 0.02<br>(0.01)    |                   | -0.00<br>(0.00)   | -0.00<br>(0.00)   |
| Trade, Hotel and Restaurants          |                   | 0.02*<br>(0.01)   | 0.02*<br>(0.01)   |                   | 0.01<br>(0.02)    | 0.01<br>(0.02)    |
| Transportation, mail and telecom      |                   | 0.01*<br>(0.01)   | 0.02*<br>(0.01)   |                   | 0.01*<br>(0.00)   | 0.01*<br>(0.00)   |
| Finance and business services         |                   | 0.02***<br>(0.01) | 0.02***<br>(0.01) |                   | 0.03<br>(0.02)    | 0.03<br>(0.02)    |
| Public services and personal services |                   | 0.04***<br>(0.01) | 0.04***<br>(0.01) |                   | 0.12***<br>(0.04) | 0.12***<br>(0.04) |
| 6-10 employees                        |                   | 0.00*<br>(0.00)   | 0.00**<br>(0.00)  |                   | 0.00<br>(0.00)    | 0.00<br>(0.00)    |
| 11-50 employees                       |                   | 0.00***<br>(0.00) | 0.01***<br>(0.00) |                   | 0.03*<br>(0.01)   | 0.03*<br>(0.01)   |
| 50+ employees                         |                   | 0.06***<br>(0.01) | 0.06***<br>(0.01) |                   | 0.04***<br>(0.01) | 0.04***<br>(0.01) |
| VA/Salary per worker, quartile 2      |                   | -0.00<br>(0.01)   | -0.00<br>(0.01)   |                   | -0.00<br>(0.02)   | -0.00<br>(0.02)   |
| VA/Salary per worker, quartile 3      |                   | -0.01**<br>(0.00) | -0.01**<br>(0.00) |                   | -0.02<br>(0.02)   | -0.02<br>(0.02)   |
| VA/Salary per worker, quartile 4      |                   | 0.01<br>(0.01)    | 0.01<br>(0.01)    |                   | -0.02<br>(0.02)   | -0.02<br>(0.02)   |
| Constant                              | 1.03***<br>(0.01) | 0.98***<br>(0.01) | 1.04***<br>(0.02) | 1.01***<br>(0.04) | 1.00***<br>(0.01) | 0.99***<br>(0.04) |
| Observations                          | 470,073           | 352,126           | 352,126           | 41,152            | 38,974            | 38,974            |

Notes: The columns show the point-estimates for  $\alpha$  from estimating the following model:  $ratio_{it} = \alpha_0 + \alpha_1 une_{it} + \beta \mathbf{X}_{it} + \varepsilon_{it}$  using ordinary least squares. Where  $ratio$  is the ratio of total vacancies to survey vacancies and  $une$  is the unemployment rate and  $\mathbf{X}$  is the set of indicators capturing firm-level characteristics. The reference group for industry is Energy and water supply. Standard errors clustered on the firm-level in parenthesis. Confidence levels are indicated as follows: \*  $p < 0.1$ , \*\*  $p < 0.05$ , and \*\*\*  $p < 0.01$ .

## 5 Concluding remarks

Using linked survey- and register-data for both Sweden and Denmark we have investigated the firm level relationship between these two commonly used vacancy measures. One is from a survey compiled by national statistical offices, while the other is register data from public employment services.

It is well known that the survey based measure covers a larger set of vacancies. Consistently with this we find that the sum of PES vacancies only constitutes around a quarter of all survey vacancies in our sample. However, the firm level structure of our data allows us to establish that the survey vacancies are not a super-set of vacancies in the PES. In particular, we find that the survey based measure must be adjusted by factor of approximately 1.2 to represent the full set of unique vacancies in these two measures.

This mis-measurement of vacancies is not necessarily a problem if it is unrelated to underlying firm characteristics and business cycle fluctuations. Shifts in the number of survey vacancies would still be informative in this case about changes in the labor market conditions. However, we show that the mis-measurement varies systematically over time and across firm characteristics. This implies that variation in survey vacancies is not only caused by changes in the actual number of vacancies, but also by changes in the level of mis-measurement.

We find that the ratio of total vacancies to survey vacancies varies from less than 1.1 to more than 1.25 over time. In other words, an observed increase in survey vacancies of 15% could solely be the outcome of changes in mis-measurement. The variation in mis-measurement over time might be a result of changes in the underlying composition of job vacancies across firm type, as the fraction of PES vacancies not included in the survey data varies from less than 60 to more than 80 percent across industries and firm size.

The degree of mis-measurement has similarities across the two countries. The number of PES vacancies constitutes about one quarter of the total number of survey vacancies in both Sweden and Denmark, but about 60% of all PES vacancies are not included in the survey data in both countries. With respect to firm level characteristics, however, the patterns varies between the two countries.

Let us conclude by suggesting some questions for future research. First, it is important to note that our findings are specific to a certain subset of firms, and not necessarily representative for the economy as a whole. It would be important to assess whether our findings carry over to a broader set of firms in the economy. Second, it would be interesting to investigate how the coefficients in the matching functions would change when using the number of unique vacancies in the survey and PES as an alternative vacancy measure.

## References

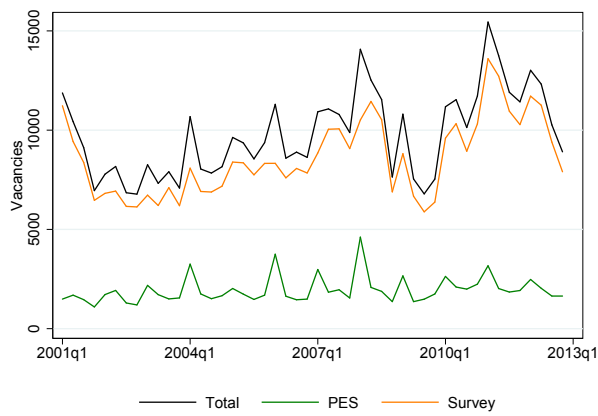
- Albæk, K. and H. Hansen (2004). The Rise in Danish Unemployment: Reallocation or Mismatch? *Oxford Bulletin of Economics and Statistics* 4.
- Berman, E. (1997). Help Wanted, Job Needed: Estimates of a Matching Function from Employment Service Data. *Journal of Labor Economics* 15(1).
- Carlsson, M., S. Eriksson, and N. Gottfries (2013). Product Market Imperfections and Employment Dynamics. *Oxford Economic Papers* 65(2).
- Danmarks Statistik (2016). Statistikdokumentation for Ledige stillinger. Technical report.
- Davis, S., R. J. Faberman, and J. Haltiwanger (2013). The Establishment-Level Behavior of Vacancies and Hiring. *The Quarterly Journal of Economics* 90, 581–622.
- Elsby, M. W. L., R. Michaels, and D. Ratner (2015). The Beveridge Curve: A Survey. *Journal of Economic Literature* 53(3), 571–630.
- Shimer, R. (2005). The Cyclical Behavior of Equilibrium and Vacancies Unemployment. *American Economic Review* 95(1), 25–49.
- Wall, H. J. and G. Zoega (2002). The British Beveridge Curve: A Tale of Ten Regions. *Federal Reserve Bank of St. Louis Working Paper Series*.
- Yashiv, E. (2000). The Determinants of Equilibrium Unemployment. *The American Economic Review* 90, 1297–1321.

## A Appendix

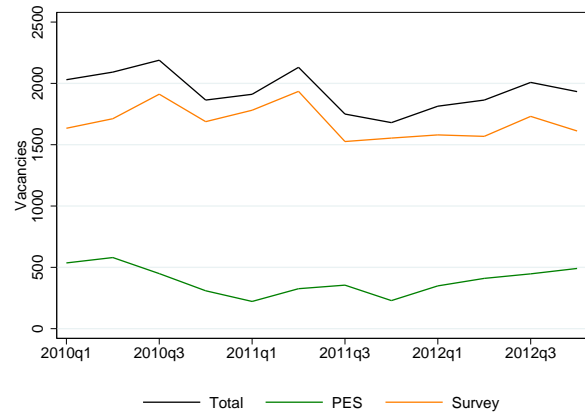
**Table A1:** Firm distribution in economy (E) and sample (S)

|   | SWE      |          | DEN      |          |
|---|----------|----------|----------|----------|
|   | <i>E</i> | <i>S</i> | <i>E</i> | <i>S</i> |
| <i>By industry</i>                      |          |          |          |          |
| Farming, fishery and mining             | 0.22     | 0.06     | 0.06     | 0.03     |
| Manufacturing                           | 0.07     | 0.20     | 0.05     | 0.15     |
| Energy and water supply                 | 0.00     | 0.01     | 0.09     | 0.14     |
| Construction                            | 0.08     | 0.08     | 0.20     | 0.29     |
| Trade, Hotel and Restaurants            | 0.17     | 0.19     | 0.07     | 0.10     |
| Transportation, mail and telecom        | 0.04     | 0.05     | 0.04     | 0.05     |
| Finance and business service            | 0.27     | 0.24     | 0.11     | 0.13     |
| Public and personal services            | 0.15     | 0.17     | 0.18     | 0.05     |
| $\chi^2$ test                           |          | 0.00     |          | 0.00     |
| <i>By firm size</i>                     |          |          |          |          |
| -5                                      | 0.57     | 0.36     | 0.72     | 0.39     |
| 5-10                                    | 0.18     | 0.17     | 0.12     | 0.14     |
| 10-50                                   | 0.20     | 0.31     | 0.13     | 0.30     |
| 50                                      | 0.05     | 0.17     | 0.02     | 0.17     |
| $\chi^2$ test                           |          | 0.00     |          | 0.00     |
| <i>By value-added/salary per worker</i> |          |          |          |          |
| Quartile 1                              | 0.26     | 0.15     | 0.25     | 0.14     |
| Quartile 2                              | 0.25     | 0.23     | 0.25     | 0.20     |
| Quartile 3                              | 0.25     | 0.29     | 0.25     | 0.29     |
| Quartile 4                              | 0.24     | 0.32     | 0.25     | 0.37     |
| $\chi^2$ test                           |          | 0.00     |          | 0.00     |

Notes: *E* : Distribution in the entire according to the Firm Database in Sweden and IDAN in Denmark, *S* : distribution in our sample, The chi-square test is a test for the two distributions being equal. For Sweden the quartiles refer to value-added per worker. For Denmark the quartiles refer to the gross salary per worker.



(a) Sweden



(b) Denmark

**Figure A1:** Aggregated total vacancies, survey vacancies and PES vacancies

**Table A2:** Dependence on missing firm-identifiers: The ratio of total vacancies to survey vacancies

|                                       | SWE               |                   |                   |                   |                   |                   |
|---------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                                       | (1)               | (2)               | (3)               | (4)               | (5)               | (6)               |
| Unemployment                          |                   |                   | -0.14<br>(0.18)   | -0.60<br>(0.49)   |                   |                   |
| Farming, fishery and mining           |                   |                   |                   |                   | 0.02**<br>(0.01)  | 0.02***<br>(0.00) |
| Manufacturing                         |                   |                   |                   |                   | 0.02<br>(0.02)    | 0.01<br>(0.01)    |
| Construction                          |                   |                   |                   |                   | 0.01<br>(0.01)    | 0.01*<br>(0.00)   |
| Trade, Hotel and Restaurants          |                   |                   |                   |                   | 0.02*<br>(0.01)   | 0.01*<br>(0.00)   |
| Transportation, mail and telecom      |                   |                   |                   |                   | 0.01*<br>(0.01)   | 0.01<br>(0.01)    |
| Finance and business services         |                   |                   |                   |                   | 0.02***<br>(0.01) | 0.03***<br>(0.01) |
| Public services and personal services |                   |                   |                   |                   | 0.04***<br>(0.01) | 0.04***<br>(0.01) |
| 6-10 employees                        |                   |                   |                   |                   | 0.00*<br>(0.00)   | 0.00***<br>(0.00) |
| 11-50 employees                       |                   |                   |                   |                   | 0.00***<br>(0.00) | 0.01***<br>(0.00) |
| 50+ employees                         |                   |                   |                   |                   | 0.06***<br>(0.01) | 0.04***<br>(0.01) |
| VA/Salary per worker, quartile 2      |                   |                   |                   |                   | -0.00<br>(0.01)   | -0.01**<br>(0.00) |
| VA/Salary per worker, quartile 3      |                   |                   |                   |                   | -0.01**<br>(0.00) | -0.01**<br>(0.01) |
| VA/Salary per worker, quartile 4      |                   |                   |                   |                   | 0.01<br>(0.01)    | -0.00<br>(0.01)   |
| Constant                              | 1.02***<br>(0.00) | 1.02***<br>(0.00) | 1.03***<br>(0.01) | 1.07***<br>(0.04) | 0.98***<br>(0.01) | 0.99***<br>(0.01) |
| Observations                          | 470073            | 124506            | 470073            | 124506            | 352126            | 104508            |
| Include pre-2009?                     | YES               | NO                | YES               | NO                | YES               | NO                |

Notes: The columns show the point-estimates for  $\alpha$  from estimating the following model:  $ratio_{it} = \alpha_0 + \alpha_1 une_{it} + \beta \mathbf{X}_{it} + \varepsilon_{it}$  using ordinary least squares. Where  $ratio$  is the ratio of total vacancies to survey vacancies and  $une$  is the unemployment rate and  $\mathbf{X}$  is the set of indicators capturing firm-level characteristics. The reference group for industry is Energy and water supply. Standard errors clustered on the firm-level in parenthesis. Confidence levels are indicated as follows: \*  $p < 0.1$ , \*\*  $p < 0.05$ , and \*\*\*  $p < 0.01$ .