Hiring, separations, vacancies and productivity

Jesper Bagger* Francois Fontaine† Manolis Galenianos‡

Ija Trapeznikova§

February 15, 2015

1 Introduction

We study empirically and theoretically the flows of workers across firms of different productivity and the hiring effort of these firms. We combine four datasets from Denmark with very detailed information on workers, firms and vacancies posted.

Using the data, we derive the patterns of hiring, separations and vacancy posting for firms of different productivity levels. We observe that some of the patterns are not fully consistent with the simple job-ladder model and propose a parsimonious theoretical model that can capture these patterns.

Section 2 presents the main patterns and describes the proposed model. Section 3 provides a detailed description of the data sources.

*Royal Holloway, University of London. Email: jesper.bagger@rhul.ac.uk
†Universit Paris 1 - Paris School of Economics. Email: francois.fontaine@univ-paris1.fr
‡Royal Holloway, University of London. Email: manolis.galenianos@gmail.com
§Royal Holloway, University of London. Email: Ija.Trapeznikova@rhul.ac.uk

2 Empirical patterns and model

2.1 Empirical patterns

We combine data from several administrative sources (see Section 3 for a detailed description). We document the patterns of hiring, separation and vacancy rates for firms of different productivity levels in the cross-section. We measure firm productivity as value added per worker.

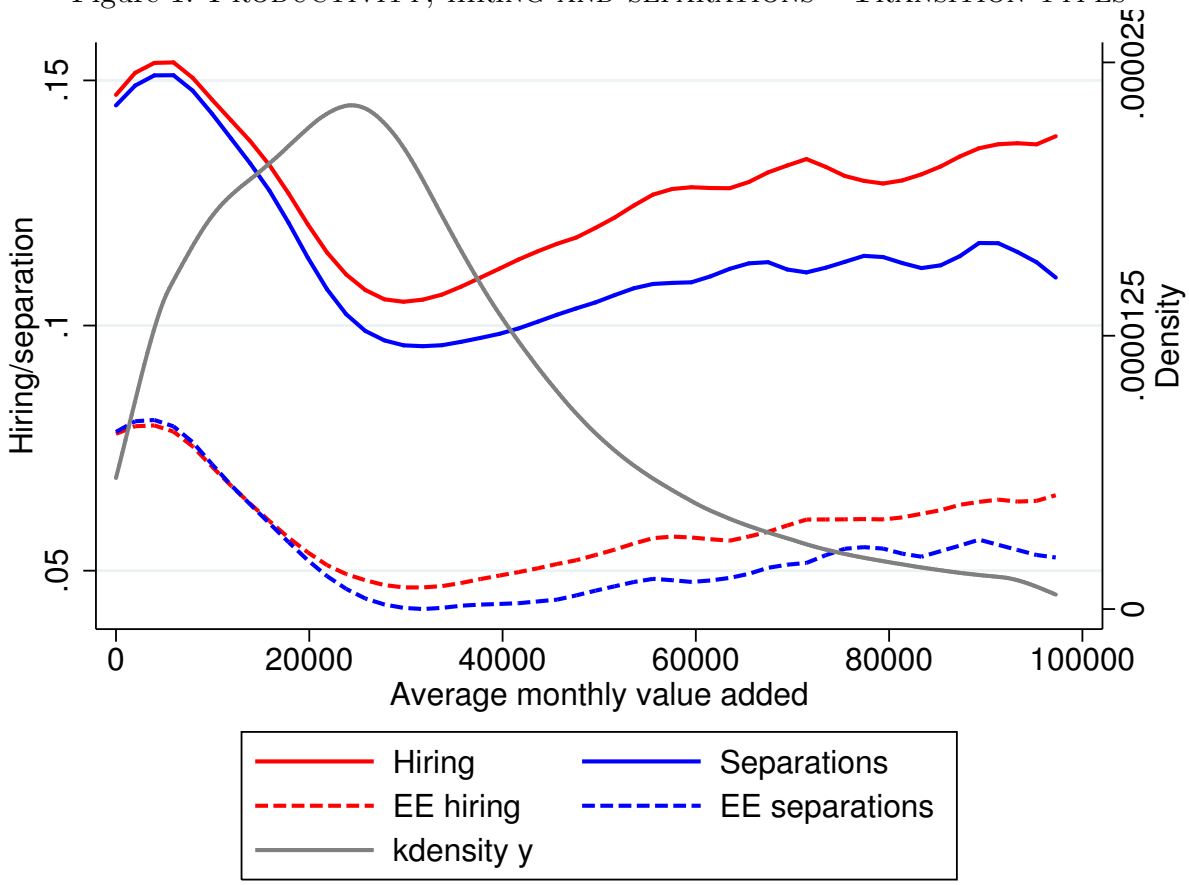
The firms' *hiring rate* has a U-shaped relation with productivity. Specifically, the hiring rate of firms at the bottom and at the top quartile of the productivity distribution is 30% and 20% higher, respectively, than that of firms at the middle two quartiles of the productivity distribution. Furthermore, approximately half of all hires are due to poaching (employment-to-employment transitions) and this proportion does not vary with firm productivity. This observation is inconsistent with the standard job-ladder model which predicts an increasing poaching rate for highly-productive firms (e.g. Burdett and Mortensen (1998) or Postel-Vinay and Robin (2002)). See Figure 1 and Table 1.

Table 1: Transitions, vacancies and yield by productivity quartile

Productivity quartile	1st	2nd	3rd	4th
Hiring rate	0.15	0.12	0.11	0.14
Percent EE hiring	51%	49%	51%	50%
Separation rate	0.14	0.11	0.10	0.11
Percent EE separation	56%	53%	53%	51%
Vacancy rate	0.19	0.18	0.17	0.30
Vacancy yield	11	7	7	6

The *separation rate* is higher for firms at the bottom quartile of the productivity distribution and is roughly constant (at a lower level) for firms at the top three quartiles of the productivity distribution. Specifically, the separation rate of firms at the top three quartiles is constant and it is 40% higher at the bottom quartile of the productivity distribution. The proportion of separations due to poaching drops modestly from 56% to 51% across the

Figure 1: Productivity, hiring and separations—Transition types



quartiles of the productivity distribution. The constant separation rate for the top 3 productivity quartiles is inconsistent with the standard job-ladder model which predicts that the separation rate declines in productivity. See Figure 1 and Table 1.

The *vacancy rate* increases in productivity and it is 2/3 higher for firms at the top quartile of the productivity distribution than at firms at the lower 3 quartiles. The vacancy yield declines monotonically in productivity and is roughly half for the top-productivity firms than the bottom ones. See Figures 2 and 3 and Table 1.

Figure 2: Productivity, hiring and separations, and vacancy rates



2.2 The model

We build a parsimonious model of random on-the-job search to account for these patterns. The model has the following fundamental features. First, the output of a match depends on firm productivity and on match quality and match quality is uncertain at the time of meeting. A signal about match quality is generated at the time of meeting and, if the match is formed, match quality is revealed at a stochastic rate. Second, match quality is more important at high productivity firms. Third, firms exert effort in hiring by posting vacancies. The surplus is split through Nash bargaining.

In this model, a firm decides how many vacancies to post and there are two reservation

Figure 3: Productivity, hiring and separations, and vacancy yield



values in a match: the reservation signal, above which a firm and a new worker form a match, and the reservation match quality, above which a worker remains at the firm after match quality is fully revealed.

High productivity firms have greater incentives to hire and post more vacancies. Since match quality is more important for high productivity firms, they also have a higher reservation signal. Therefore, they have a higher vacancy rate and a lower vacancy yield than lower-productivity firms.

At low productivity firms a large proportion of contacts with applicants leads to a hire which results in a high hiring rate despite low levels of vacancy posting. At high productivity firms, the opposite holds: a small proportion of contacts with applicants leads to a hire

but many vacancies are posted, leading again to a high hiring rate. At firms with medium productivity, the hiring rates are lower.

A firm's separation rate consist of quits (when workers are poached by more desirable firms) and layoffs because match quality is revealed to be below reservation, in addition to a constant job destruction rate. Highly productive firms have low quits but high layoff rates because their reservation match quality is high (this holds under standard assumptions on the distribution of match quality). Therefore, their separation rate due to job-to-job transition is lower than that of lower-productivity firms but their overall separation rate is at a similar level. Low productivity firms, on the other hand, have high quit rates but low layoff rates.

3 Data

This section describes four main data sources used in the analysis, as well as the construction of variables of interest. The end product is a comprehensive panel of Danish firms including virtually all private businesses in the economy for 2003-2007.

The firm level panel is build from a matched employer-employee dataset. Persons are identified by their CPR-number. All individuals residing legally in Denmark are registered in the CPR register (in Danish: "Centrale Persons Register", abbreviated CPR) with a unique CPR number. The CPR-number is the sole identifier of an individual in relation to the Danish authorities.

Firms are business entities or enterprises, carrying out one or more activities at one or more locations. Throughout the paper we identify firms using their CVR-number, obtained from the Central Business Register (in Danish: "Centrale Virksomheds Register", abbreviated CVR). This register contains primary data on all businesses with economic activity in Denmark, regardless of economic and organizational structure. A firm's CVR-number is the standard way businesses identify themselves to stakeholders, for example when issuing invoices. Some of the data we use is collected under a different firm identifier, denoted the

SE-number.[1] The relationship between CVR- and SE-numbers is one-to-many (but most often one-to-one), and information registered under SE-numbers can be aggregated to the level of the CVR-number. Appendix A provide a more detailed description of the two different firm identifiers.

3.1 Payroll data

We obtain employment information from MIA data (Monthly Data Reports of A-Income), made up of mandatory payroll data for all firms with employment in Denmark. The data is being collected to settle tax-at-source payments and is reported to the Danish Central Customs and Tax Administration by employers.[2] MIA is mostly payroll data covering the period May, 1997 to December, 2007. There is no information on which days within a given month a given worker was employed, nor the hours worked, and nor the amount paid. The unit of observation in the MIA data is a employer-employee-match month. The worker is identified by his or her CPR-number. The employer is identified by an SE-number.[3] We aggregate employment data to the level of CVR-numbers.[4]

For each firm-month we now record the number of hires and separations. Let N_{jt} be the number of recorded in firm j during period t. Let H_{jt} be the number of worker hired by firm j during month t. Finally, let S_{jt} be the number of workers separated from firm j during month t. The nature of the data is such that N_{jt} is the number of workers who received a payment from firm j during month t. A month-t hired worker is a worker on the payroll in month t, but not on the payroll in month $t-1$. A month-t separated worker is a worker who was on the payroll in month $t-1$, but not on the payroll in month t. That is, firm-level

[1] A CVR-number identifies legal business entities. An SE-number identifies fiscal business entities. The notion of a fiscal business entity is mostly used in relation to tax payments.

[2] In Denmark, labor income is either denoted A-income or B-income. A-income is taxed at the source, whereas B-income is is not. A typical example of B-income is a payment for freelance work. The employment data in MIA covers all workers receiving A-income. Virtually all labor income is A-income.

[3] As the MIA data is used for tax purposes it is generally of very high quality and with virtually no measurement error. However, we have identified a problem with the MIA data relating to April, 2007.

[4] We thank Henning Bunzel for providing this mapping.

monthly employment N_{jt} has the following law of motion:

$$N_{jt} \doteq N_{j,t-1} + H_{jt} - S_{jt}. \tag{1}$$

As it turns out, out preferred measurement of hires and separations, adds an additional stability condition. Specifically, we define a *permanent hired worker* as a worker who was hired into a job that will last at least three months. Similarly, define a *permanent separated worker* as a worker who is separating from a job that has lasted at least three months. Let PH_{jt} denote firm-level permanent hires, and PS_{jt} denote permanent separations. Obviously, $PH_{jt} \leq H_{jt}$ and $PS_{jt} \leq S_{jt}$. Conversely, define *temporary hired workers* and *temporary separated workers* as $TH_{jt} \doteq H_{jt} - PH_{jt}$ and $TS_{jt} \doteq S_{jt} - PS_{jt}$. The law of motion for firm-level monthly employment is now written as

$$N_{jt} \doteq N_{j,t-1} + PH_{jt} + TH_{jt} - PS_{jt} - TS_{jt}. \tag{2}$$

Finally, the employment histories from MIA allow us to identify job-to-job transitions. A job-to-job transition (at month t) occurs when a worker is hired by firm j in month t , was employed by another firm j' in month $t-1$, and the firm-j' employment relation ship ends no later than in month $t+2$, as measured in the MIA data. Hence, we allow for some overlap in payments.[5] Let EEH_{jt} be the number of workers hired by firm j through job-to-job transitions in month t, and let EES_{jt} be the number of workers separated from firm j through job-to-job transition in month t.

From the payroll data we now construct a firm-level panel containing information on monthly employment N_{jt}, hires H_{jt}, permanent hires PH_{jt}, hires through job-to-job transitions EEH_{jt}, separations S_{jt}, permanent separations PH_{jt}, and separations through job-

[5]This definition of a job-to-job transition may include some job-to-nonemployment-to-job transitions. Indeed, a worker may be laid off from firm-j' on the first day of month $t-1$ and find employment at firm j on the last day of month t. The payroll data is monthly, so the worker will receive a payment from firm j' in month $t-1$ and a payment from firm j in month t, and we will record a job-to-job transition.

to-job transitions EES_{jt}. The panel covers virtually all firms with employment in Denmark over the period 2003-2007.

3.2 Background information on firms

Background information on firms are obtained from IDA, the main Danish matched employer-employee data. This data is constructed form a host of administrative records by Statistics Denmark and is available for the full period of focus here, 2003-2007. We use the IDA FIRM-files that contains information on all firms (identified by their CVR-numbers) with economic activity. From the FIRM-files in IDA, we are able to identify whether a firm is a public sector firm or not, and also obtain industry indicators on the NACE 1.1 format. After some basic cleaning and consistency checks on the FIRM files, we merge the FIRM data to the firm-level employment panel. 94% of all employment panel firm-months are merged to FIRM data. The share of FIRM-observations matched to MIA data is equally high at 92%.

3.3 Firm productivity

We shall measure firm productivity by value added per worker. Information on value added is obtained from value added tax (VAT) data. This information is obtained from MOMM-files, also held by Statistics Denmark. It provides firm-level information on purchases and sales of all VAT-liable businesses in Denmark during the relevant period, 2003-2007. In Denmark, a business enterprise must register for VAT if its annual turnover is expected to exceed 50,000 DKK. The VAT declaration frequency depends on the firm's annual turnover. Firms register VAT information at the tax authorities monthly if their turnover exceeds DKK 15 mill., quarterly if their turnover is between DKK 1 mill. and DKK 15 mill., and every six months if turnover is below DKK 1 mill. The reporting frequency for a given firm is also provided in the MOMM data.

Using VAT reports it is straightforward to back out sales and expenditures on interme-

diates for each firm. Different firms will have difference reporting frequencies. In this paper, where the focus is on cross sectional heterogeneity, it suffices to look at annual data. For each firm-year we therefore aggregate sales and expenditure information to an annual level. With a slight abuse of notation, let t index years. Then, let R_{jt} and M_{jt} be firm j's turnover and expenditures (in intermediates) in year t. Firm-level annual value added is

$$Y_{jt} = R_{jt} - M_{jt}. \tag{3}$$

Merging the annual value added panel we match 83% of the employment panel firm-months to value added data. The share of observations in the value added panel matched to the employment panel is much lower. Presumably this is due to many firms in the value added data having no employees.

3.4 Vacancy data

The most innovative feature of the data used in this paper is a panel of firm-level vacancy data. This data stands apart from the other data sources described above in that it is not sourced from public registers. Instead, it is obtained from a major private Danish online job board. The data made available to us covers the period June 1st, 2002 to December 31st, 2009. We shall use the period Jan 1st, 2003 to December 31st, 2007.

The vacancy data of course included vacancies posted on this specific job board. However, the company in question also operates a sophisticated search engine which daily scans the Danish part of the World Wide Web for online job posting. This includes ads in online newspapers, on individual firm's web pages, other job boards, public job centers, etc. The job board portal operated by this company is therefore "the place to look" for jobs in Denmark, and a conservative estimate is that it covers more than 90% of the vacancies posted online in Denmark during the relevant period 2003-2007.[6] Crucially, the job board that collects

[6]Source: Personal communication with the director of the data provider.

this data also operates a sophisticated algorithm to detect identical ads posted at multiple (online) outlets.

The data made available to us contains, at a daily frequency, new vacancy postings. Hence, we measure the vacancy inflow.[7] The data is rich. Indeed, when extracting online ads from other online outlets, the company's search engine effectively retains all text in the ad. As it turns out, a large fraction, about 60%, of all the ads contains the firm's CVR-number. This is because, as explained above, the CVR-number is the main administrative identifier for a firm vis-a-vis its stakeholders. For example, the CVR-number is often embedded in the company logo. Using this information, we are therefore able to merge vacancy information *at the firm level* to the existing stock of Danish register data, as described above. We are not aware of any other panel of vacancy postings at the firm-level, and certainly none that can be matched to comprehensive administrative records on both employers and employees.[8]

For the purpose of this paper, we have extracted the following information from the raw (daily) vacancy data: Posting date, CVR-number (when available), and occupation. [9]. We aggregate vacancy postings to a monthly level (denote the number of posted vacancies by firm j in month t by V_{jt}), and merge the resulting monthly firm-level vacancy posting panel to the firm-level employment panel. 81% of the firm-months in the vacancy panel with non-missing CVR-number are matched to a firm-month in the employment panel. Of course, the share of firm-months in the employment panels that is matched to vacancy data is much smaller because most firms do not post vacancies in most months, and about 40% of the vacancy observations do not include information on the CVR-number of the posting firm.[10]

[7] A vacancy is defined to be new if it is less than three weeks old when located.

[8] Vacancy data from the same source is used by Brodersen, Dimova, and Rosholm (2014). However, these authors do not have information on the CVR-number of the posting firm, and have an altogether very different focus.

[9] Incorporating occupation information in the analysis is still to be completed.

[10] It is of interest to establish whether the firms with missing CVR-numbers are different from those with CVR-numbers. At this stage, we have checked whether the distribution of occupations differ between vacancy postings with and without CVR-number information. It does not.

3.5 Analysis panel

The constructed firm-level panel is comprehensive. It contains information on employment, hiring, separations, productivity, and vacancy posting. Employment, hiring and separations, and vacancy postings are available at a monthly level. Following Davis, Haltiwanger, and Schuh (1996) we construct firm-level hiring rates (h_{jt}, ph_{jt}, and eeh_{jt} for all hires, permanent hires and job-to-job hires, respectively), separation-rates (s_{jt}, ps_{jt}, and ees_{jt} for all separations, permanent separations and job-to-job separations, respectively) and vacancy rates (v_{jt}) as

$$x_{jt} = \frac{X_{jt}}{(N_{jt} + N_{j,t-1})/2} \in [0, 2], \tag{4}$$

for $X_{jt} = H_{jt}, PH_{jt}, EEH_{jt}, S_{jt}, PS_{jt}, EES_{jt}, V_{jt}$. The constructed rates range between 0 and 2, with 2 indicating either a closing firm, $S_{jt} = N_{j,t-1}$ and $H_{jt} = 0$ so $N_{jt} = 0$, or a starting firm, $H_{jt} = N_{jt}$ and $N_{j,t-1} = 0$ so $S_{jt} = 0$.[11] As our focus in on cross sectional heterogeneity, we then average the monthly firm-level observations to an annual level. Hence, the analysis panel contains annual information on the average monthly employment, and hiring, separation and vacancy rates, as well as labor productivity as measured by average monthly value added per worker.

We conduct a minimum amount of data trimming prior to the analysis. First, we discard the top and bottom 1% of the annual cross sectional distribution of value added per worker. Second, we discard the top 1% of the firm size distribution. Second, we discard all with more than 150 workers. This is very few firms as the 99th percentile in the firm size distribution is 93. Third, we discard firm-years where the firm is either starting up or closing down.[12]

Basic descriptive statistics on this is provided in the first column in Table 3.5. After trimming, the annual panel contains 387,910 observations (firm-years). Of these, 37,259, or about 10%, are associated with firm-yes with vacancy postings. For the results presented in

[11] We refer to Davis, Haltiwanger, and Schuh (1996) for further details and statistical justification for measuring growth rates in this way.

[12] All our empirical findings are robust to the trimming.

this version of the paper we rely entirely on permanent hires and separations. Hence, in what follows hiring rates and separation rates refers to permanent hiring and separation rates, as defined in (2). The average monthly hiring rate is 0.12 while the separation rate is 0.11. The vacancy rate is 0.20 and the vacancy yield is 6.72.[13] Finally, we find that the correlation between productivity, as measured by value added per worker, and firm size, as measured by the size of the workforce is only weakly correlated. This is consistent with findings in Lentz and Mortensen (2008), using Danish data.

Table 2: Descriptive statistics

Number of observations	387,910
Average hiring rate	0.12
Average separation rate	0.11
Number of observations with vacancy data	37,259
Average vacancy rate	0.20
Average vacancy yield	6.72

References

Brodersen, S., S. Dimova, and M. Rosholm (2014): "Analyzing How ALMPa affect the Demand Side of the Labor Market," Discussion paper, Aarhus University.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996): *Job Creation and Destruction*. MIT Press, Cambridge and London.

Lentz, R., and D. T. Mortensen (2008): "An Empirical Model of Growth Through Product Innovation," *Econometrica*, 76(6), 1317–1373.

[13]The reported monthly hiring and separation rates are high, albeit consistent with findings in Trapeznikova (2014) who find that hiring and separation rates in Denmark are about twice that typically found in the US. In other words, the Danish labor market is highly mobile.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Trapeznikova, I. (2014): "Employment Adjustment and Labor Utilization," Discussion paper, Royal Holloway, University of London.

APPENDIX

A Details on the data

A.1 Person and firm identifiers

Persons are identified by their CPR-number. All individuals residing legally in Denmark are registered in the CPR register (in Danish: "Centrale Persons Register", abbreviated CPR) with a unique CPR number. The CPR-number is the sole identifier of an individual in relation to the Danish authorities.

There are two main ways of identifying business entities or enterprises. A business entity carries out one or more activities at one or more locations, and, as we shall describe below, is often a single legal entity. Hence, a business entity is best described as firm, not an establishment. The Central Business Register (in Danish: "Centrale Virksomheds Register", abbreviated CVR), established in October 1999, is the central register containing primary data on all businesses with economic activity in Denmark, regardless of economic and organizational structure. CVR covers both public and private businesses. A firm's CVR-number is the standard way businesses identify themselves to customers and public authorities, for example when issuing invoices.

The SE-number identifies a business in Stamregistret for Erhvervsdrivende (the SE-register), a register established in 1985. Denmark operates a Value Added Tax (VAT) system, a tax-at-source system where employers withhold the income tax of its employees, as well as a system of mandatory pensions contributions. The main function of SE-register is to identify businesses vis-a-vis the tax authorities when settling VAT, income tax payments, or tax payments for self-employed persons, and pension contributions.

The relationship between CVR-numbers and SE-numbers is not straightforward. A CVR-number identifies a legal entity vis-a-vis customers, authorities, and other stakeholders. The SE-number represents a fiscal business entity vis-a-vis the tax authorities. When the CVR-number was introduced in October, 1999, most legal business entities were assigned their pre-existing SE-number as a CVR-number. Similarly, almost all entering businesses appear with a single SE-number per CVR-number. As a business evolves, grows or shrinks it may obtain additional SE-numbers, or may shed some SE-numbers, under the same CVR-number. Of course, a business entity may also split up into multiple legal business entities, thus obtaining multiple CVR-numbers, or may close down some legal business identities, abandoning CVR-numbers. Hence, most businesses consist of one CVR-number and one SE-number, however some legal business entities have multiple SE-numbers. An SE-number is associated with one and only one CVR-number at a given point in time.

Some of the data we use are collected by CVR-number (e.g. vacancy data) and some are collected by SE-number (e.g. payroll data, VAT data, and income data). We are able to map SE-numbers to CVR-numbers, and, since the relationship between CVR- and SE-numbers is one-to-many (although very often one-to-one), we conduct our analysis at the level of CVR-numbers.