Economics
Working Papers

2018-3

The Danish Matched Employer-Employee Data

Emmanuele Bobbio and Henning Bunzel

Abstract:
The Danish data constitutes a unique source of information: it covers and links together the universe of persons, establishments and firms for more than 30 years. In addition workers' histories are constructed at the weekly level. The data has the potential to become a benchmark in social research, but access has been limited by the lack of systematic documentation. The chapter provides an introduction, with a particular emphasis on those portions more commonly used in labor market research. We establish a wide variety of descriptive statistics that can serve as a reference for future studies. The second part is dedicated to a more detailed analysis of wages and flows, for which the data is uniquely well suited. Returns to seniority are low on average and even lower for workers coming from unemployment; wage changes between jobs are large, but often negative and in these cases the wage remains persistently lower. Labor market turnover is high, the unconditional job and unemployment hazard display strong negative duration dependence. Also, higher wages are associated with a lower probability and duration of unemployment. Finally, we use the data to revisit the recent debate on the "ins and outs" of unemployment: the reduced form decomposition assigns 2/3 of unemployment volatility to the ue rate; the ue rate is procyclical, while the eu rate countercyclical and leads unemployment; the ee rate is procyclical, so the separation rate is essentially acyclical. These results are remarkably consistent with findings from U.S. data.

AARHUS BSS
DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

The Danish Matched Employer-Employee Data

Emmanuele Bobbio and Henning Bunzel

Abstract The Danish data constitutes a unique source of information: it covers and links together the universe of persons, establishments and firms for more than 30 years. In addition workers' histories are constructed at the weekly level. The data has the potential to become a benchmark in social research, but access has been limited by the lack of systematic documentation. The chapter provides an introduction, with a particular emphasis on those portions more commonly used in labor market research. We establish a wide variety of descriptive statistics that can serve as a reference for future studies. The second part is dedicated to a more detailed analysis of wages and flows, for which the data is uniquely well suited. Returns to seniority are low on average and even lower for workers coming from unemployment; wage changes between jobs are large, but often negative and in these cases the wage remains persistently lower. Labor market turnover is high, the unconditional job and unemployment hazard display strong negative duration dependence. Also, higher wages are associated with a lower probability and duration of unemployment. Finally, we use the data to revisit the recent debate on the "ins and outs" of unemployment: the reduced form decomposition assigns 2/3 of unemployment volatility to the *ue* rate; the *ue* rate is procyclical, while the *eu* rate countercyclical and leads unemployment; the *ee* rate is procyclical, so the separation rate is essentially acyclical. These results are remarkably consistent with findings from U.S. data.

Key words: Danish MEE data, wage dispersion, job flows, personal wage dynamics, turnover, ins and outs of unemployment
JEL codes: J21, J31, Y1

Emmanuele Bobbio
Bank of Italy, via Nazionale 91, IT-00184 Rome, e-mail: ecbobbio@gmail.com, phone: +390647921.

Henning Bunzel, corresponding author
Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, e-mail: hbunzel@econ.au.dk, phone: +4587165517.

1 Introduction

The Danish *MEE, Matched Employer-Employee Data* is constructed by using administrative records from *Statistics Denmark* (hereafter, DS). It has two distinctive features: the extensive coverage – the universe of *persons* (*workers*) and *firms* for more than 30 years between 1980 and *2013* – and the possibility of linking together records from different portions of the data – e.g. *employees* at the same *employer* or *members* of the same *household*. In addition, *LMDG, Labor Market and Dynamic Growth* at Aarhus University has constructed *spells* with weekly frequency for each *person* in the population between 1985 and *2013*. Together, these registers provide a reliable and detailed source of information portraying the socio-economic development of the country over an extended period of time. The data has the potential to become a laboratory for social research and serve as a benchmark for a variety of studies. Unfortunately the lack of systematic documentation has limited access to it.

This chapter aims at filling this gap, with particular attention to the portion of the data more commonly used in labor market research, and requiring employer-employee matched data. [1]. It describes *populations* and *identifiers* and the overall structure of the data in Section 2 and Section 3. Then, the content of some registers is explained in more detail, Section 4.[2] The second part establishes some stylized facts on *wages* Section 6, and *flows* Section 7. The Danish data is uniquely well suited for this task, because of the panel dimension and the availability of detailed worker histories. In the process we derive descriptive statistics for the population of *persons*, *establishments*, and *firms* which may serve as reference for future studies, Section 5.

Wage dispersion is low, while dispersion in value added per worker is much larger, suggesting that firm heterogeneity is important. Next, we turn to person wage dynamics. Returns to seniority are low or nil, especially in the second part of the life-cycle. Unemployed workers who find a job face an even flatter wage stream. Instead, wage changes between jobs are large, but often negative; when the wage declines it remains persistently low.

Regarding flows, labor market turnover is high. High wages correlate with a lower unemployment probability and, perhaps surprisingly, with a higher probability of transitioning to another job. Also a higher wage in the previous or following job is associated with a higher unemployment hazard . These facts suggest that worker heterogeneity plays an important role in the labor market. Then, we use the data to revisit the recent discussion on the cyclical behavior of the "ins and outs" of unemployment – Shimer (2005, 2007), Hall (2005), Fujita and Ramey (2009), Elsby

[1] The chapter discusses the version of the data available through LMDG/ECONAU, at Aarhus University. The LMDG version of *IDA, Integrated Danish Labour Market data base* includes persons aged 15-74, while the original data covers all persons. The paper uses documentation found on the web pages of Statistics Denmark, http://www.dst.dk, or more specifically http://www.dst.dk/da/TilSalg/Forskningsservice/Dokumentation. A large part of this documentation has been created by the "KOR Højkvalitetsdata" project

[2] This part is more technical and meant for prospective users of the data. And more detailed information is available in Bunzel (2016).

et al. (2009). Results are remarkably close to findings in those studies: the decomposition assigns 2/3 of unemployment volatility to changes in the *eu rate* – though this number should not be given a structural interpretation; the *ue rate* is strongly procyclical while the *eu* rate is strongly countercyclical and leads unemployment by 1-2 quarters. However the *ee rate* is procyclical, resulting in an acyclical job-finding rate (the flow rate to non-participation is acyclical)

To conclude this section, previous work on the Danish *MEE data* has focused on wage dispersion. Rosholm and Svarer (2004), Christensen et al. (2005) are early contributions, while Mortensen (2003) summarizes (and advances) the literature with several applications to the data. More recently, Bagger et al. (2014) exploits the availability of education and labor market experience to account for standard Mincer regressors in the Postel-Vinay and Robin (2002) framework. Bagger et al. (2013) and Bagger et al. (2016) focus on the joint distribution of productivity and wages. Taber and Vejlin (2011) allow for compensating differentials and Bobbio (2010) for replacement search on the side of *firms* to reconcile wage dynamics, wage inequality and labor market mobility. Finally Bagger and Lentz (2014) exploit the availability of workers' wage and employers' productivity to evaluate sorting.

2 Overview

2.1 Data structure

Table 1 summarizes the portion of the Danish data discussed in the chapter. We use uppercase, italic letters for registers and uppercase slanted letters for variable names. The data can be thought of as being structured in two components: *IDA*, *Integrated Danish Labour Market data base* and the *FS*, *Firm Statistics*.

IDA is a longitudinal data set covering the universe of *persons* , *establishments* and all *primary jobs* existing in the last week of November. It extends from 1980 to 2013 and has annual frequency. *IDA* consists of five data sets:

- *IDAP, IDA, Persons, labour market attachment and experience*
- *IDAN, IDA employments/jobs*
- *IDAS, IDA establishments*
- *IDFI, IDA Income tax reporting units*
- *IDAPALL, IDAP merged with demographic, income and educational data*

On the other hand, the purpose of *FS* is to monitor *economic activity* at the *firm level*. For 1992 and up to 1998 *FIGT, Old firm statistics, Employment and Economic Variables* and *FIGF, Old firm statistics, accounts data.* are constructed using *VAT* and *custom forms* and it covers only some industries: *manufacturing*, *construction*, *retail*, and from 1998, *wholesale* and *mining*. In 1999 the data is extended to the universe of *firms* – including *public firms*. This portion is properly referred to as the *FIRM, Firm statistics* but we will use the acronym *FS* for both parts, unless the

Table 1 Danish employer-employee matched data, synoptic table

Data sets		Identifiers[a]					Key	Time	Freq.	Population	Remarks
		PNR	*ARB-GNR*	*ARB-STK*[b]	*LBNR*	*CVRNR*					
IDA	*IDAP persons*	×					*PNR*	80-13	year	Persons age 15-74 living in Denmark end of year	Labor market variables, income, year of entry and exit, experience, unemployment, unemployment insurance.
	IDAPALL persons	×					*PNR*	80-13	year	As *persons*	End of year covariates, e.g.: gender, education, family status, kids, partner and parents' ID, address, worker category, experience, income sources, assets, liabilities
	IDAN jobs	×	×		×		*PNR, TYPE, ARBGNR*[c]	80-13	year	Jobs in the last week of November	Variables describing job, e.g. job classification, private and public sector, establishment ATP pension payments, full-time - part-time.
	IDAS establishments		×	×	×		*LBNR* or *ARBGNR ARB_STK*	80-13	year	Establishments, at least 1 employee at some point during the year	End of year covariates, e.g.: workforce, worker turnover, wage-bill, industry, company type, municipality
FS	*FIRE accounts*					×	*CVRNR*	95-13	year	Firms in: *construction, retail, manufacturing, mining*. From 1998 also: *wholesale*. From 1999 all firms	Accounting statistics based on surveys and tax reports. A weighted sample: all firms $if \geq$ 50 FTE, 50%$if \geq 20FTE$, 20%$if \geq 10FTE$, 10%$if \geq 5FTE$
	FIRM firms					×	*CVRNR*	99-13	year	All economic active firms	Firm statistics based on *FIRE*, *MOMS*, *EBS*.General firm covariates like number of employees, industrial sector, net capital value,import, export, wage, purchase of intermediate products, sales, value added
	FIDA	×				×	*PNR, SSTILL*[d] *PSTILL,*	95-13	year	Primary and secondary jobs in last week of November at firms in *FS*	Links a person to a firm in FIRM or FIRE. The firm identifiers of these data sets do not correspond to firm identifiers in *IDA*.
	SPELL	×	×	×			*PNR DELSTART, DELSLUT*[e]	85-13	week	As *persons*	Workers' labor market histories; it includes employer's ID

Overview of the registers discussed in the chapter: identifiers, key, population, frequency and brief description. [a] Identifiers are encrypted and used as keys in one or more data sets. A key value uniquely identifies an observation. [b] From 2008 ARBNR is used instead of ARB_STK. [c] The primary job in the last week of November is identified by TYPE equal to H,A,S,M [d] Type of primary job, zero if the record corresponds to a secondary job [e] Beginning and ending week of a spell.

distinction is necessary. Frequency is annual and the *firm identifier* in these data may change from one year to the next for the same firm, so the data is not longitudinal. [3] *FS* provides a rich set of accounting measures, *FIRE, Accounts statistics*. It also includes the *UHDI, Discretional Foreign Trade data* – not discussed here – with information on *values*, *quantities*, *country of origin*, and *country of destination* at the eight digit *commodity* level.

The universe of *employer* differs between *IDA* and *FIRM* but *DS* has constructed a register, *FIDA, Key between IDA persons and firm statistics* , linking the November cross-section of jobs to *FS*.

Finally *DS* makes available several other administrative registers which are used by *LMDG* to construct *spells*. The data covers all persons in *IDA* for every week between 1985 and *2013*. Covariates for *employees* and *employers* can be merged from *IDA* and *FS* – see Table 1 for an indication of how the various registers can be merged together.

2.2 Populations and identifiers

Different registers may describe different portions of a population; frequency and time span may vary as well. However, the data defines a consistent set of identifiers allowing the linking of information about a given entity from various sources, or the relating of different entities to one another – e.g. members of a same *family* , *employees* at an *establishment* , *establishments* within a *firm* . For more information, see Bunzel (2016).

Persons: Each *person* in the Danish population is assigned a social security number at birth or, for a non citizen, when the person migrates to Denmark for the first time. The number is scrambled and stored in the variable *PNR, Person identifier, CPR number* it consistently identifies the person over time and across registers.

Establishments: An *establishment* is a well defined physical entity located at a particular address, and producing one or a few goods. There exist two equivalent identifiers for *establishments*. The first is the pair: *ARBGNR* and *ARB_STK*. Here, *ARBGNR* designates the *fiscal identifier* (as we discuss below) while *ARB_STK* specifies an *establishment* within this *ARBGNR*. [4] This pair is related to the *fiscal identifier* which may change from one year to the next. Instead, a second identifier, *LBNR, Establishment identifer which is permanent over time* is constructed by DS with the purpose of identifying an *establishment* across years. The *establishment* is considered to be the same if one of three criteria is met: same owner and industry; same owner and workforce; same workforce and either same address or same industry.

[3] Currently, work is carried out to create a *firm identifier* which is consistent over the years.

[4] If the *firm* has only one *establishment*, *ARB_STK* is missing until 1994 and equal to 1 after this year. In some DS data sets the variable is named *DSKOD*. *DSKOD* has been renamed to *ARB_STK* in ECONAU data sets.

IDA relates *employers* identified by a fiscal identity, and *employees* at *establishment* level. However, certain jobs cannot be associated to a particular *establishment* but only to a particular *employer*, for example: sailors, sales representatives, or *employees* working from home. For these jobs DS introduces the concept of a *fictitious workplace*. [5] In this chapter the term *establishment* designates an actual production unit, while a *workplace* may be a production unit or a *fictitious workplace*.

Firm Identifiers: The *firm identifier* and indeed the concept of a *firm* as an *economic unit* varies among registers and over time.

The point of departure is a *legal firm unit* which is a unit responsible for contracts with other firms and reporting and paying taxes to the government. The legal unit can be a privately owned firm. The owner is named *private owner* if it has no employees, otherwise *employers*. Or the firm may be incorporated, i.e. a company.

A commonly used identifier is *SENR, A firm account number for tax and custom payments*, which is used when paying taxes and tariffs or reporting income, trade, etc. to the government. *SENR* has existed in all years.

In 1999 *CVR, Central Firm Register* was introduced and all legal units are assigned a unique identifier *CVRNR*.

A firm may have many *SENR*, and in special cases many firms use the same *SENR*. Hence, there is not a one-to-one relationship between *SENR* and a *firm*.

The major part of small firms has only one *SENR* which is equal to its *CVRNR*. Some *economic units* have a *SENR*, but not a *CVRNR*. The structure of a *SENR* and a *CVRNR* is similar. Hence, only the source can identify it as a *SENR* or a *CVRNR*. Before 1999 *DS* had created *JURNR, Legal firm or organization id* which for private firms would be a *SENR*. [6]

Economic unit or firm: The *economic unit* is the smallest set of legal economic firms which can make economic decisions like marketing, investments, and hirings [7]. Most firms consist of just one legal firm unit. *CVRNR* is used as identifier in *FS*. *FS* covers the universe of economic units in Denmark.

The statistical unit in *IDA* is the *establishment* rather than the *firm*. However, *IDA* does have a concept of a firm which is identified by the *fiscal identifier* on which employment and salaries have been reported, *ARBGNR, IDA identifier for unit reporting wages for taxation*. *ARBGNR* is not consistent over time because an *economic unit* can choose to get a new *CVRNR* or *SENR*.

The *IDA* data includes *ARBGNR8*which is equal to a *SENR* or *CVRNR* for private firms 1980-2007 and gradually for public firms from 1999-2007. From 2008 and forth *CVRNR* is used, but it must still be interpreted as a *fiscal identifier*.

---

[5] For a job at a fictitious workplace *LBNR* is 0000000000. Before 1991 the largest *establishment* within that *firm* was used for unknown/fictitious workplace. Starting in 1991 *ARB_STK* is assigned a string whose first character is a letter, or a number between 8869 and 8900.

[6] In current LMDG data sets *JURNR* has been renamed *CVRNR* because they are being used to identify firms in firm and trade statistics.

[7] A group of economic firms may form an *economic group* but this entity is not covered in this chapter

It is important to note that a *firm* in *FS* is not the same as a *firm* in *IDA* and therefore the sets of *establishments* belonging to the two firm definitions differ. *FIDA* includes primary and secondary jobs from *IDAN* together with the *CVRNR* of the *firm* in *FIRM* or *FIRE* in which the worker is employed.

Today *DS* uses the concepts *economic unit* and *economic group* which is a group of economic units. The legal units forming Economic Units are identified by *CVRNR*. In this chapter *firm* is used instead of *economic unit*, and one of the *CVRNR* associated with *economic unit* is used to identify the firm in the data sets.[8]

3 The registers

3.1 IDA

IDA covers the universe of *persons* residing in Denmark at the end of the year and all *establishments* public and private, having at least one employee at some point during the year. Also, it covers the cross section of *primary jobs* and *secondary jobs* existing in the last week of November. *IDA* extends from 1980 to *2013* with annual frequency. The data is organized in four data sets, each for a different entity: *IDAP, IDA, Persons, labour market attachment and experience*, *IDAS, IDA establishments* and *IDAN, IDA employments/jobs*, *IDFI, IDA Income tax reporting units*. The universe of *persons* each year is derived from a static version of the *CPR, Central Population Register*, see Persons and Family, Section 3.4

LMDG has constructed *IDAPALL* which merges *IDAP* with commonly used variables from several other registers.

3.1.1 The IDAP and IDAPALL registers

The population is formed at the end of the year. For each person the data set gathers end-of-year information based on several administrative registers. *IDAP* contains yearly data on the association with and the yearly activity and outcome on the labor market, see Table 2. *IDAPALL* contains information on *demographics* and *education*: gender, age, marriage, country of birth, highest completed and ongoing *education*, see Table 3. On *family structure*: partnership arrangement, partner, number and age of kids, parents, see Table 4. And finally on *welfare* and *income and taxes*. See Table 5. Unemployment, Employment and Education are stock sampled, while all other information is flow sampled. Note that the data allows one to link observations for members belonging to the same *family* e.g. partners, parents and kids.

[8] New data sets have a variable *OK_NR* which identifies an *economic unit*. It is not a *CVRNR*, but can be mapped to one or more *CVRNR*.

Table 2 *IDAP* selected variables for LABOR MARKET

Variable	Description
AKASSE	*Code for unemployment fund*
ARLEDGR	*Annual rate of unemployment*
ATPAR	*Number of years as employee*
ERHVER	*Job experience from 1980 in 1000*
EXPYEAR	*Number of years with labor market experience. An improved version of ERHVER.*
JOB_LOEN_BELOEB	*Sum of salary and wages for a job*
PENSALD	*Age of retirement, public sector*
POTLEAR	*Number of years since 1980 with unemployment or being unemployment insured*
SENAFAR	*Most recent exit from labor market*
SENSTAR	*Most recent entry into labor market*
STARTAR	*First year in the labor market*
SUMGRAD	*Sum of rate of unemployment since 1980*)

Table 3 *IDAPALL* selected variables for DEMOGRAPHICS and EDUCATION

Variable	Description
KOEN	*Gender, 1=Male, 2=Female, 9=Unknown*
ALDER	*Age, ultimo year*
OPR_LAND	*Country of origin*
FOED_DAG	*Date of birth extracted from CPRNR.*
IE_TYPE	*Immigrants, descendants, persons of Danish origin*
HFAUDD	*Highest completed education*
UDD	*Code of an education or educational activity*
ALMAUDD	*Generel education*
ALM_VFRA	*Time of completed general education*
ERHAUDD	*Vocational education*
ERH_VFRA	*Highest completed vocational education*
HF_VFRA	*Time of highest completed education*
HFAUDD	*Highest completed education*
HFFSP	*Highest completed education, main group*

Table 4 *IDAPALL* selected variables for FAMILY and GEOGRAPHICS

Variable	Description
AEGTE_NR	*CPRNR, person ID of spouse*
FAELLE_NR	*CPRNR of spouse or cohabitant*
ARB_KOM	*Municipality of location of the establishment*
BKOM	*Municipality of residence*
FAMILIE_ID	*FamilyID, a unique identifier of the family. Break in 2007.*
FAR_NR	*Person ID, father*
MOR_NR	*Person ID of mother*

3.1.2 The establishments' register IDAS

The *establishments*' register is constructed using yearly taxable income reports. As such, it covers any *establishment*, public or private, employing at least one worker at some point during the year – even if the *establishment* opens and/or shuts down during the year. The register only includes physical *establishments*. *IDAS* contains data about the dynamics of the establishment over time, various measures of total employment and the wage-bill, along with information on ownership, private or public;

Table 5 *IDAPALL* selected variables for INCOME AND TAX

Variable	Description
ARB_SEKTOR_PO	*Sector for establishment, private/public*
BRUTTO	*Gross income*
DISCOALLE_INDK	*DISCO for primary employment*
DISPON_NY	*Income after taxes and interests (Total income with deduction of interest expenses*
ERHVERVSINDK	*Income from own business*
FORM	*Capital*
FORMUEINDKNY	*Capital income, 1980*
HELTID_DELTID_KODE	*The total work extent of a person grouped into full-time or part-time exployment ultimo November.*
KOEJD	*Sales value of property in Denmark per 31.12.*
LOENINDK	*Wage and salaries*
NETOVSKUD	*Surplus from own business*
PERSAMLINKNETRENT_NY	*Total income incl. net interest payments*
PERSBRC	*Industry of person's workplace or establishment*
PERSONINDK	*Total personal income without capital income*
QAKTIVFNY_05	*Assets ultimo year excl. Pensions*
QPASSIV	*Total of registered liabilities*
QPASSIVN	*Liabilities according to bank reports*
QPENSIALT	*Total pensions*
QRENTUD2	
QSPLINDK	*Taxable income*
SLON	*Total salary and wages from annual tax report register*
SOCIO	*Socioeconomic classification*
SOCIO02	*Socioeconomic status, 2002*
STIP	*Stipends from the State Education Fund*
TIMELON	*Imputed hourly earnings of employment*
TIMELON_VL	*Imputed hourly earnings of employment. Improved version of TIMELON.*
TLONKVAL	*Indicator for quality of the variable TIMELON*

Table 6 *IDAS* selected variables

Variable	Description
ARB_ANSAT_NOV	*Number of employees in primary job, TYPE=H, last workday in November*
ARB_ANSAT_SUM	*Number of persons employed during the year at the establishment*
ARB_AARSV_SUM	*Number of full-time equivalent workers during the year*
ARB_LOEN_SUM	*Yearly sum of salary and wages for all employments at the establishment.*
ARB_KOM	*Municipality of location of the establishment*
ARB_SEKTOR	*Sector for the establishment.*
ARB_BRANCHE_DB07	*Industry code for establishment (db07)*

also it provides the industry code and the identifier for the municipality where the *establishment* is located. – see Table 6.

3.1.3 The jobs register, IDAN

Similarly to the *establishments* register, the jobs register is constructed from tax withholding records and, accordingly, it covers employer-employee relations. However, it differs from the other components of the Danish data in that it mainly describes the cross-section of primary and secondary jobs observed in the last week of November. A worker can hold multiple jobs at a given time, and the register has some information about all jobs. On the *employer* side each record is linked to the

Table 7 *IDAN*, selected variables

Variable	Description
ARBGNR	*IDA identifier for unit reporting wages for taxation*
ARBGNR8	*IDA identifier for unit reporting wages for taxation in form of JURNR, SENR,* or *CVRNR.*
CVRNR	*CVRNR, CVRNR, Legal unit identifier*
LBNR	*Establishment identifer which is permanent over time*
STILL	*Position on labor market*31-33=manager 34=white collar 35=skilled 36-37=unskilled worker. Coding changes in 1996.
TILKNYT	*Extent of employment for primary job.*
TYPE	*IDA: Type of employment for a job in IDAN.DREAM: A person is an immigrant (I), descendant (E) or Dane (D).*
JOB_LOEN_BELOEB	*Sum of salary and wages for a job*
JOB_TYPE	*Description of number of hours in the primary and secondary job offered by the firm.*
TIMELON	*Imputed hourly earnings of employment*
TIMELON_VL	*Imputed hourly earnings of employment. Improved version of TIMELON.*
TLONKVAL	*Indicator for quality of the variable TIMELON*

establishment identifier, *LBNR* or just *ARBGNR*, if the worker cannot be associated to a specific *establishment*.

The register conveys information on earnings and hours. It contains the total gross compensation received by the employee, including bonuses, and *ATP Payments*. The *ATP Payments* depends on the type of contract, on a four point grid reflecting hours/week. DS exploits the information to compute the *hourly wage*. The measure is noisy, but an indication of its quality is given for each record. Table 7 describes these variables.

3.2 FIRM and FIRE

FIRM, Firm statistics and *FIRE, Accounts statistics* complement *IDA* in two respects. First, the appropriate unit of analysis in *IDA* is the *establishment* rather than the *firm*. In IDA, the *firm* is defined as a fiscal entity, which may or may not be an *economic unit*. Instead, *FIRM* focuses on the firm as an *economic unit*. Second, while *IDA* covers *employers demographics* *FIRM* and *FIRE* provide a detailed picture of firms' *economic activity*.

The Firm Statistics are constructed using *VAT* and *custom forms* . As such, they cover only some industries: in 1992-1997 *manufacturing* *construction* and *retail* in 1998 *wholesale* and *mining* are added. From 1999 *firm* covers all public and private firms with a basic set of economic activity variables. *DS* introduces a survey allowing one to monitor *firms* which are not subject to *VAT* or customs duties – e.g. the financial industry and the public sector.

The accounting register, *FIRE, Accounts statistics*, has a large number of accounting variables, the same sector coverage as *FIRM*, but is based on detailed surveys of all firms operating in Denmark, possibly supplemented with tax information. The survey is administered to a rotating sample representing companies with 5 *employees* or more; those with fewer than 5 *employees* account for approximately 14% of employment – see Section 4. Sampling weights are as follows: all *firms* with more

Table 8 Accounting statistics (*FIRE*), selected variables

Variable	Description
OMS	*Revenue*
AUER	*Goods produced for own consumption or investment.*
ADR	*Other operating income*
DLG	*Change of stock*
KRH	*Purchase of raw material, intermediate material, finished goods and packaging (excl purchase*
KENE	*Purchase of energy*
KLOE	*Purchase of sub-contracters and salaries (non-employed)*
OEEU, *ANEU*, *EKUD*[a]	*Other external expenses*, *Other external expenses*, *External expenses (apart from items of secondary character)*
UDHL	*Expenses for rent*
ULOL	*Expenses for longterm lease and operation leasing*
UDVB	*Costs for lease of manpower from other firm*
UASI	*Expenses for purchase of minor capital goods etc*
SEUD	*Secondary expenses*
LGAG	*Wages and salaries*
PUDG	*Pension expenses*
AUDG	*Other expenses for social security*
AARE	*Result of the year (after corporate tax)*
ATIT	*Increase, total*
AFAT	*Investments, annual exit*
AT	*Assets, total*
AAT	*Property, plant and equipment totally*
JKOD	*Code for the source of the firm's accounts reports*

Variables are grouped in categories: measures of output, variable costs, labor costs, investments/assets. [a] for 95-98, 99-01 and since 2002, respectively.

than 50 *employees* (in FTE) or having a "large" turnover; 50% of *firms* with 20 to 49 *employees* ; 20% of *firms* with 10 to 19 *employees* and finally 10% of *firms* with 5 to 9 *employees* . The resulting sample includes $9,000$ companies out of a population of approximately $190,000$; however, it covers 2/3 of employment because the average *employer* is large, approximately 1000 workers – see Section 4.[9] DS verifies the information against several administrative registers and the result is deemed very reliable.

FIRE gathers information from these various sources which are not covered by the *VAT* and custom duty register, nor by the survey, have values imputed based on a firm labor share (in FTE) of sampled firms within the same sector. The variable *jkod* indicates the source and quality of the information used to construct the record; in particular, it allows one to select only those records whose information is obtained from the survey, so that appropriate sampling weights can be applied.

Table 8 displays selected variables for the accounting register. They can be used to construct measures of *economic activity* which are commonly used in empirical work. The table is organized as follows. The variables in the first part of the table quantify the output of the firm, while those in the second and third reflect the variable cost and the labor cost, respectively. Value added can be computed as output

[9] This is the size of the *employer* averaging *firms* across workers, i.e. average *firm* size using the *firm*'s employment share as weight.

minus variable cost. Net-investment is gross investments minus investment disposals.

3.3 Linking FIRM and *IDA* , *FIDA*

DS links jobs in the November cross-section to *FIRM, Firm statistics* thus matching the *employer* definitions from the two portions of the Danish data, *IDA* and *FIRM*. You merge employees to *firm* by first merging f.inst. *IDAN* and *FIDA* on *PNR* then merging *FIDA* and *firm* on *CVRNR* from *FIDA*.

FIDA contains both primary and secondary jobs. If you only want primary jobs you can condition on *PSTILL* not being missing. *IDA* workplaces are merged to firms by using *LBNR* in *IDAN*. You may also use the LMDG data set *MEE, Matched Employer-Employee Data.*

3.4 Persons, Families and Household

Data sets have been created for the three units*persons*, *families*, and *households*from *PSD, Statistics Denmark, Population Statistics Database.*

The definition of Families and Household was changed in 2008 from *C families* to *E families.* The *C family* is defined as parents and children, 18 years or less, living with the parents. The *E family* includes children, 25 years or less, living with the parents. Different variables are used to describe *C families* and *E families.*

These data stes are generated yearly from *PSD. PSD* is being updated daily from the administrative register *CPR, Central Population Register* including updates concerning previous periods. However, the yearly data sets generated from *PSD* for statistical purposes are not updated.

The population of *PSD* data sets are persons living in Denmark January 1st. However, *IDA* data sets are based on persons living in Denmark December 31st, hence the *PSD* data sets are lagged one year when forming the *IDA* compatible data sets *PERSONER*, *FAMILIE*, and *HUSSTANDE*.

In order to correspond to the *IDA*population, these data sets use the *C family* data sets for the years 1980-2007, and the *E family* data sets from 2008 and onwards. After the data is lagged one year the periods are 1979-2006, 2007-*current year.*

The *ECONAU*, datasets, 1980-*current year*, contain common variables for both definitions, but notice that the definition may differ for the periods 1980-2006 and 2007-*current year* see the description of variables in *PERSONER*, *FAMILIE* and *HUSSTANDE* in Table 9, 10, and 11.

Table 9 Selected variables from *PERSONER*

Variable	Description
BKOM	*Municipality of residence*
POSTNUMMER	*Postal code is a four-digit number connected to a geographical area.*
ALDER	*Age, ultimo year*
KOEN	*Gender, 1=Male, 2=Female, 9=Unknown*
FOED_DAG	*Date of birth extracted from CPRNR.*
FAR_NR	*Person ID, father*
MOR_NR	*Person ID of mother*
FM_MARK	*Indicates if this person lives with one or both parents.*
CIVST	*A person's civil status*
CIV_FRA	*Date of civil status*
FAMILIE_ID	*FamilyID, a unique identifier of the family. Break in 2007.*
FAELLE_NR	*CPRNR of spouse or cohabitant*
C_STATUS	*A person's status in the C_FAMILY to which he/she is a member or constitutes*
C_TYPE	*Type of C family*
PLADS	*A person's status in the E-family.*
STATSB	*Citizenship*
VAN_VTIL	*Date of arrival of immigrants*

Table 10 Selected variables from *FAMILIE*

Variable	Description
AAR	*The year of observation*
ALDAELDST	*Age of oldest child in family. Break in 2007.*
ALDYNGST	*Age of youngest child in family. Break in 2007.*
ANTBOERNF	*Number of children in family. Break in 2007*
ANTPERSF	*Number of persons in family. Break in 2007.*
BKOM	*Municipality of residence*
C_STATUS	*A person's status in the C_FAMILY to which he/she is a member or constitutes*
C_TYPE	*Type of C family*
FAMILIE_ID	*FamilyID, a unique identifier of the family. Break in 2007.*
FAMILIE_TYPE	*Family type. Break in 2007.*
HOVEDPERSON	*Head of family. Break in 2007.*
PAPNR	*Partner of head of family in FAFA and FAM, see FAELLE_NR.*

Table 11 Selected variables from *HUSSTANDE*

Variable	Description
ALDAELDST	*Age of oldest child in family. Break in 2007.*
ALDYNGST	*Age of youngest child in family. Break in 2007.*
ANTBOERNH	*Number of children in family. Break in 2007*
ANTEFAM	*Number of efamilies in HUST, see ANTFAM*
ANTPERSH	*Number of persons in family. Break in 2007.*
BKOM	*Municipality of residence*
BOPIKOM	*Address in municipality(road, house number, house character, floor, side or door number)*
HUSTYPE	*Type of household. Break in 2007.*

3.5 Workers' labor market history, SPELL, Labour market spells

IDA contains detailed information on labor market outcomes. However, frequency is annual and values are cumulated over the year – for example with regard to the number of weeks of unemployment. *DS* provides administrative registers covering taxes, *welfare* and education. These registers specify the starting and ending date - week or month - for each record – for example the day a person registers as unemployed and the last day s/he receives benefits. *LMDG* at Aarhus University uses this information to construct labor market histories for all persons in *IDA*. [10] For these persons the register records the labor market state each week from 1985 to 2003, together with an identifier of the *establishment*(*ARBGNR8*, *ARB_STK*) for employed workers. Then, covariates of *employer* and *employee* can be merged from the registers described above.

The register is constructed by successively re-assigning a state to a worker in each week, starting with the information that is considered less reliable. Table 12 lists and describes each state in the order it is assigned.[11,12] Spells may overlap, if the worker holds more than one job at the same *firm* and a portion of the history may be missing if the worker is out of the country for an extended period of time.

New Spell Data. The new spell data set cover the period 1985-2014. [13] The construction consists of three steps.

1 Construction of **basic spells** with similar structure for the states

- Persons living in Denmark
- Employment
- Unemployment
- Public benefits
- Education
- Pensions
- Other outside labor market

The spells may overlap both within and across spell types. The spells have a common set of identifiers *PNR*, *SPELL_CVRNR*, and *SPELL_LBNR*, which are also included in all other data sets that can be merged with the spells.

---

[10] This is a continuation of the work done by the labor market group at the Aarhus School of Business.

[11] Rehabilitation is a form of public support for persons with severe conditions, such as mental disability.

[12] More precisely, weeks are numbered from 3 to 992, the November cross-section corresponding to weeks: 49, 101, 153, 206, 258, 310, 362, 414, 467, 519, 571, 623, 675, 727, 780, 832, 884, 936, 988. The information is organized by spell, or by year if the spell extends across years. Each record contains the beginning (*delstart*) and ending week (*delslut*). The variable *tilstand* stores the string of states assigned, from right to left (the leftmost being the most reliable). Finally, *foranst* provides information on the assignment before the last one, which can be used to revert to that state.

[13] See Bunzel and Hejlesen (2016).

Table 12 Labor market states in *SPELL, Labour market spells*

State	Description	Aggregation
N[1]	Residual state, the worker may be out of the labor force	N
F[2]	Unemployed, vacation	Previous state
A[2]	Unemployed, vacation paid by unemployment found	Previous state
S[3]	Self-employed	S
E[3]	Employed	E
G[4]	In active labor market program	U
X[5]	Miscellaneous, see *foranst*	E/S/N/R[a]
B[6]	Parental leave	E/S/N[a]
Y[6]	On (paid) leave, due to health reasons	E/S/N[a]
R[6]	Rehabilitation (e.g. mental illness)	R/N
H[6]	Early retirement	R
P[6]	Receiving retirement pension	R
O[6]	Receiving elderly pension	R
K[7]	Cash benefits	U
T[7]	Temporarily unemployed, receiving benefits (e.g., seasonal jobs)	U
U[7]	Receiving unemployment benefits	U

The order reflects the assignment, from less (*1*) to more (*7*) reliable. The last column summarizes the aggregation procedure which reduces the number of states to employment (E), self-employment (S), unemployment (U), non-participation (N) and retirement (R). [a] See footnote 15

Table 13 States in New Spells

E	Employed
U	Unemployed
C	In education
R	Retired receiving pension
N	Residual state, out of labor force

Spells with *economic activity* also contain *SPELL_TOTAL_SALARY* and *SPELL_TOTAL_HOURS*.

2 Constructing a **primary spell** for each state from any overlapping spells within each group. The **primary spell** will have links to the basic spells used for its construction.
3 Constructing coherent, non-overlapping spells for each person in the period 1985-2014. For employment this is done at the firm level and at the establishment level.

The new spells have fewer states, see Table 13 but each spell may have links to subspells which can be used to refine the state, for example for Employed into specific labor market programs.

4 Descriptive statistics

4.1 Preparing the data

We remove from the *SPELL, Labour market spells* register persons with missing or overlapping history spells (6 and 0 persons respectively). Then, we reduce the list of states to employment (E), self-employment (S), unemployment (U), non-participation (N) and retirement (R) – we refer to "inactivity" as the sum of non-participation and retirement.[14] The aggregation procedure is summarized in the last column of Table 12: temporary unemployment and participation in an active labor market program are treated as unemployment; all forms of retirement are pooled together and rehabilitation is re-coded as non-participation; sick and parental leave are treated as self-employment or employment, if information on the *employer*'s identity is available, or non-participation. [15] Finally, A and F are overwritten with the previous state – or the record *ARBGNR8*is discarded, if it is the first for that worker. [16]

Job spells are defined at the level of the *employer* legal identifier. This may result in spurious job-to-job transitions due to changes in the *firm*'s fiscal code, an issue which is addressed at establishment level using the time-consistent *LBNR, Establishment identifer which is permanent over time* constructed by DS (10.8% of transitions). [17] Next, we join two employment spells if a worker is observed returning to the previous *employer* within 13 weeks, and we re-code a non-participation spell lasting less than a year as unemployment. [18]

In the second stage we merge firm and employee covariates into the spell data. We use *IDAN* November job to import the wage from the jobs register. Next, *FIDA, Key between IDA persons and firm statistics* is used to relate the *ARBGNR8*from *IDA* with *CVRNR* from *firm* hence allowing us to import *establishment demographics* from *IDAS* and *firm* accounting statistics from *FIRE*.

Below we compute descriptive statistics for the data and then focus on the private sector for a more detailed analysis of wages and flows. [19] It should be kept in mind that the distinction between non-participation and unemployment becomes arbitrary

---

[14] We wish to thank Jesper Bagger and Rune Vejlin for sharing with us the details of the procedure.

[15] In particular, if the first character in *tilstand* is either B or Y we check the second character; if it S, or if it is E and the *employer* identifier is available (*ARBGNR*), then the state is coded as self-employment or employment, respectively; otherwise, it is coded as non-participation, which in practice is a residual state. The same procedure is applied to X if *foranst* is either 106, 108, 109 or 209; instead, if it is either 104, 105, or 143 the state is coded as self-employment, and if it is either 118, 119, or 121 as retirement.

[16] For example, F may correspond to the case of an unemployed person who had accumulated days of vacation while at the previous *employer* .

[17] Suppose that *ARBGNR* changes across two consecutive records, while *lbnr* remains the same: While the change in the *firm* identifier suggests a change in job, the worker is still employed at the same *establishment* .

[18] In Denmark, a worker is entitled to take a leave from work for up to three months.

[19] Similar descriptive statistics for the new spell data is available from Bertheau (2016).

when the worker is not registered as unemployed. Also, the distinction between job-to-job transitions and lay-offs is problematic, because the worker may receive a notice and find a new job before losing the current one; similarly, a worker moving to a different *employer* may be inactive for some time.

4.2 Descriptive statistics

Table 14 displays descriptive statistics for the data, every five years and for the last year of the period. The number of entities in each year is reported in the bottom part and is roughly constant over time. An person is observed for 14.3 consecutive years on average, an *establishment* for 6.5 years and a *firm* for 6.0 years, if its is defined as a legal entity (*ARBGNR, IDA identifier for unit reporting wages for taxation*), or 3.34 years, if it is defined as an economic entity (*CVRNR, CVRNR, Legal unit identifier*, the series runs from 1999 onward). Figures are sensible, though males are slightly over-represented.

Several trends in demographics can be discerned.[20] Education has been growing over time with a shift towards college attendance and, to a lesser extent, vocational training. The fraction of households with cohabitating partners has increased while the number of households with kids has dropped slightly, reflecting a fertility change which, however, has been mild relative to southern Europe. The fraction of the population age 15-74 participating in the labor market has remained roughly constant throughout the period (the increase in the last year of the sample is due to a spurious rise in unemployment, as discussed below). However, the percentage of persons attending school has increased, as well as the fraction of retirees – possibly as a result of early retirement policies implemented during the 80s. The gap between the male and the female participation rate (not reported) ranges between 7 and 8 percentage points over the period.

Regarding the composition of the labor force, private employment has been growing at the expense of the public sector and self-employment. Unemployment is hump-shaped (excluding the last column). This may reflect business cycle movements, particularly the 1988 and 1993 recessions; or, it may be the consequence of changes in labor market institutions. [21] The same pattern can be found in DS official figures, which, however, are substantially lower (by 5 to 6 percentage points). This depends on the particular procedure chosen to aggregate labor market states: it treats government sponsored employment as unemployment; and, more importantly, residual-state spells lasting less than a year are recoded as unemployment – while

[20] The trend in experience is due to the fact that the measure is constructed starting with 1963, so it under estimates experience for any cohort born prior to 1948.

[21] Denmark traditionally favors employment flexibility and decentralized negotiations, while providing support through generous unemployment benefits and through active labor market programs. The approach dates back to the end of the 19th century, but has been formalized in the 20th century, between the 80s and the beginning of the 90s.

Table 14 Employer-employee matched data, descriptive statistics

	1985	1990	1995	2000	2003
male fraction	50.36	50.37	50.43	50.40	50.32
age	40.15	40.39	40.83	41.79	42.36
experience[a]	8.08	9.74	11.20	12.74	13.56
primary edu.	48.20	44.58	39.85	34.97	32.98
high school	6.45	7.05	8.26	8.79	8.61
vocational	30.68	32.32	33.43	34.51	34.57
some univ.	2.93	3.17	3.64	4.32	4.72
univ. graduate	8.72	9.47	10.71	12.41	13.47
postgraduate	3.02	3.40	4.10	4.99	5.65
married	41.69	39.22	37.89	38.12	37.71
cohabitating	9.08	10.25	11.03	11.55	11.48
living with kids	29.01	27.15	26.09	26.27	26.68
married l.w. kids	20.84	18.17	16.68	16.70	16.72
labor force	67.62	67.39	66.11	66.62	68.26
retired	13.00	13.77	15.20	15.77	16.41
in school	10.08	10.40	10.98	11.83	11.99
other non-participation	9.30	8.45	7.71	5.78	3.34
Labor force composition:					
private employment	42.51	45.64	45.92	51.05	48.13
self-employment	9.06	7.67	6.84	6.38	5.55
public sector	30.11	29.48	26.99	25.61	24.82
unemployment	13.63	15.11	16.66	12.10	17.14
$<$ 52 weeks	1.55	1.35	1.51	1.36	9.94
$<$ 104 weeks	7.35	4.06	4.41	3.72	2.91
T_e[b]	118.2*	108.0*	103.2*	72.3	nr
T_{ee}	265.8	233.8	170.7	nr	nr
T_{eu}	392.3	nr	nr	nr	nr
T_u	41.14*	79.32*	64.20*	55.33*	nr
T_{ue}	50.84*	110.83*	90.41*	69.22	nr
hourly wage, DKK[c]	92.35	128.70	144.58	170.34	187.25
assets (in 1000)	—	—	304.18	481.97	581.93
liabilities (in 1000)	177.33	150.88	161.29	—	—

The persons register and the jobs, *establishments* and accounting registers are merged with the *SPELL, Labour market spells* register, 1985-2003. [a] Available from 1963, i.e. valid for cohort 1948 onward [b] Mean duration (Kaplan-Meier), decomposition by ending state under competing risk assumption: employment, unemployment, non-participation (not reported); $*$: 90^{th} percentile of corresponding survival function is defined; estimates are not reported (nr) when third quartile is undefined. [c] Mean, Danish Krone (DKK)

a portion corresponds to actual non-participation. This aspect of the aggregation procedure also explains the (spurious) increase in the last year of the sample.

Next, the table displays the mean duration of a job spell and of an unemployment spell (Kaplan-Meier estimates).[22] Also it shows figures by ending state, estimated under the competing risk assumption. The mean survival time is underestimated due to censoring of the largest observation and we only report figures such that the third quartile of the corresponding survival functions are defined (a "star" indicates that the 90^{th} percentile is defined as well). Labor market turnover appears to be high and comparable to U.S. levels, with the important caveat of long term unemployment,

[22] Transitions to retirement and to non-participation are considered as "events".

Table 14 (continued)

	1985	1990	1995	2000	2003
unskilled[d]	22.29	22.15	22.22	11.54	6.17
clerical	29.82	29.67	26.59	16.52	15.88
skilled	12.77	12.48	13.35	37.54	42.55
manager	27.33	28.24	27.26	23.09	16.56
Employer's descriptives					
manufacturing	21.69	20.94	20.53	18.74	17.05
construction	6.68	6.06	5.89	6.26	6.18
trade	14.47	14.31	15.12	15.54	15.60
transport	4.96	7.42	7.05	7.14	6.82
finance/real-estate	9.55	10.28	10.29	11.79	12.25
K[e]	—	—	—	448.41	523.80
VA	—	—	—	307.01	344.96
W	—	—	—	721.31	971.61
Firm's descriptives		2000		2003	
FTE < 5[f]		71.93	14.78	71.41	14.06
< 10		11.95	7.85	11.76	7.37
< 20		8.25	10.68	8.86	10.93
< 50		5.13	14.64	5.28	14.37
< 100		1.45	9.45	1.44	9.04
< 250		0.86	12.18	0.82	11.36
< 500		0.24	7.63	0.25	7.77
<1000		0.13	8.28	0.11	7.02
≥1000		0.06	14.51	0.07	18.09
VA,W (Spearman correlation)[g]		71.02		68.55	
,labor share		−62.61		−61.64	
,FTE		5.38		5.35	
,K		25.62		24.19	
FTE,W		19.83		19.96	
,K		−1.55		1.72	
K,W		4.59		3.80	
N. observations	1985	1990	1995	2000	2003
persons	3786282	3859594	3873241	3825667	3789188
ESTABLISHMENTS	172565	170636	163597	166910	166385
firms (*ARBGNR, IDA identifier for unit reporting wages for taxation*)	136990	132093	124599	127917	127284
firms (*CVRNR, CVRNR, Legal unit identifier*)				126658	126995
with accounting data				8993	9113

[d] Break in the coding of underlying variable in 1996 [e] Mean fixed assets, value added and wage bill per *FTE, full time employment* in 1000 DKK, FTE weighted; FTE=full time equivalent employees, 37 hrs.×48 weeks [f] Fraction of firms and – in parenthesis – employment share by class size [g] VA, W, K, FTE defined as above, not FTE weighted

which may be due to the generous welfare system. Assuming for simplicity that transitions are generated by independent and homogeneous Poisson processes, the contribution of job-to-job transitions is approximately 50%, that of transitions to unemployment 25%, and that of non-participation 15% – T_{ej}/T_e. [23]

The second part of the table displays figures for the occupational composition, which should be interpreted with caution due to a structural break in the coding of

[23] These figures are comparable to estimates for the U.S. by Bjelland et al. (2011), who find job-to-job transitions to be twice as likely as transitions to unemployment.

the underlying variable that occurred in 1996. As for the sectoral composition, the economy has gradually moved away from agriculture (not reported) and *manufacturing* and towards services, especially finance and real estate.

The mean nominal wage has approximately doubled between 1985 and 2003, but the GDP deflator (computed from chain linked series) has increased by 53.2%, reducing real wage growth to 32.2% over 18 years (or a compounded 1.57% a year, on average). Based on the *firm level* accounting data, the aggregate labor share, averaging between 1999 and 2003, is .69 and the aggregate capital output ratio is 1.89, on an annual base.

Looking at correlations, more productive *firms* pay higher wages, but a smaller fraction of value added, resulting in a concave relationship between productivity and firm average wage – productivity being defined as value added per full time equivalent employee (FTE). The link between productivity and size (in FTE) is weak. More capital is associated with higher value added, but capital intensity does not strongly correlate with the wage. In conclusion, only value added per worker and size correlate with the wage.

On average, an *establishment* and a *firm* employ 12.05 and 15.65 workers respectively, in FTE units, and a *firm* has 1.30 *establishments* . [24] The firm size distribution is highly skewed: $3/4$ of *firms* employ fewer than 5 *employees* and account for 14% of employment only, while the 0.4% of *firms* employing at least 250 *employees* account for approximately 30% of employment.[25] The co-worker mean is 313 at the *establishment* level and 3688 at the *firm level*.

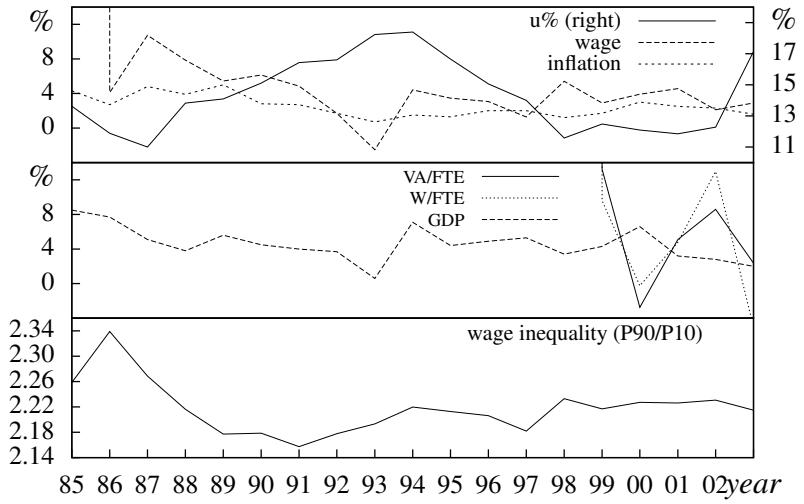
5 Wages

Figure 1 summarizes the evolution of aggregate *prices* and *quantities* over the period. The top panel displays the unemployment rate together with the percentage change of the mean wage from one year to the next, and of the GDP deflator; in the middle panel we report the growth rate of the mean value added and the mean wage bill per FTE – FTE weighted – as well as nominal output growth.[26] Finally, the last panel displays the $90/10$ percentile ratio of the the wage distribution, as a measure of wage inequality. As mentioned above, Denmark has experienced two recessions, in 1988 and 1993, with an intervening period of slow growth. At the same time unemployment has increased, while inflation and most notably wages have de-

[24] Average *establishment* size has grown over time from a minimum of 11.17 in 1985 to a maximum of 12.87 in 2002. The number of *establishments* per *firm* has increased as well, though not as systematically, but the pattern may be due to changes in fiscal practices affecting the definition of a *firm*. With this caveat, average firm size has also increased from a minimum of 14.07 in 1985 to a maximum of 16.81 in 2002. When considering the definition of a *firm* as an *economic unit* (*CVRNR, CVRNR, Legal unit identifier*) numbers are comparable.

[25] Note that the firm size distribution satisfies the Pareto principle, the 80% of smaller *firms* employing approximately 20% of workers.

[26] Nominal GDP and its deflator are computed by DS and available through Eurostat.



Top panel: unemployment rate (right scale), percentage change of the mean nominal wage and of the GDP deflator (source: Eurostat). Middle panel: mean (FTE weighted) of nominal value added and wage bill per FTE, percentage change, and nominal GDP growth (source: Eurostat). Bottom panel: $90/10$ percentile ratio of the hourly wage.

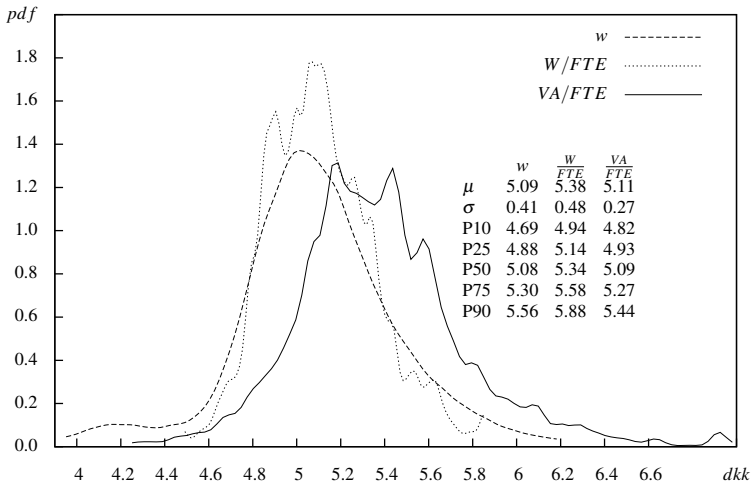
Fig. 1 Unemployment, output, wages and inflation

clined. In the second part of the period the economy has grown at a sustained rate. Inequality has remained low, though increasing.

Figure 2 displays the wage distribution in the last week of November 2000, together with the distribution of the wage bill and value added per FTE . Quantities are in logs. The densities are non-parametric estimates using the wage and *firm* covariates of each employee observed in the private sector – firm *quantities* are scaled from the annual to the hourly level using a factor of 48 weeks $\times$ 37 hours/week. The idea for the figure is drawn from Mortensen (2003) and, in particular, Bagger et al. (2013) and Bagger et al. (2016), who attempt to explain the joint distribution of output and wages as the result of frictions and *firm* and worker heterogeneity.

Several facts are worth noting. First, wage inequality is low, with the $90/10$ percentile ratio at 2.39. Indeed, the wage distribution is well approximated by a log-normal – Mortensen (2003). Second, wage dispersion at the employer level is lower, suggesting that within firm heterogeneity is significant. Finally, dispersion in value added per FTE is larger than workers' wage dispersion, hinting at the importance of firm heterogeneity – Lentz and Mortensen (2008).[27] Structural work on wage dispersion stresses the role of job mobility, as proposed in the seminal paper of Burdett

[27] At the sectoral level, *construction* and transport, storage and communication exhibit the lowest degree of dispersion in the wage bill and value added per *FTE, full time employment* , respectively, while finance and real estate displays the highest degree of dispersion along both dimensions. In finance and real estate and in *construction* the two variables are similarly dispersed relative to



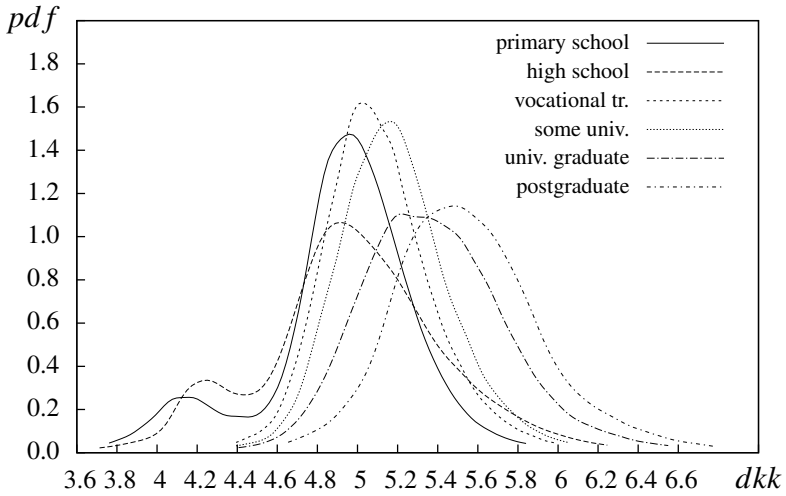
Private sector, worker's wage and FTE weighted firm's value added and wage bill per FTE in logs. Non-parametric estimates: normal kernel, rule of thumb. Results displayed on the 1-99 percentile range.

Fig. 2 Productivity and wages, Nov. 2000

and Mortensen (1998). Mortensen (2003) summarizes and advances this line of research using the Danish data – see also Rosholm and Svarer (2004) and Christensen et al. (2005). The framework has strong implications for wage dynamics over the employment spell. Wages should rise with seniority, and unemployment should reset the process – a fall from the "wage ladder." The seminal paper by Postel-Vinay and Robin (2002) on the French data further develops the framework. More recently, work by Bagger et al. (2014) extends this line of work by exploiting the availability of education and experience in the Danish data to disentangle the role of human capital accumulation in the early part of a worker's career. Given its length and breadth the Danish data is ideal to address these sorts of questions. Furthermore, as argued above, the Danish labor market appears remarkably fluid, the mean job duration for the entire private sector being 120 weeks. We compare the wage distribution and the wage profile of workers with different education levels, and in the next section we analyze worker flows in and out of unemployment and between jobs. The sample is stratified by highest completed education according to the following six categories: primary school, high school, vocational training, short length university education, university graduate and, finally, postgraduate education.

---

one another, while dispersion in value added per FTE is substantially higher than dispersion in the wage bill per FTE in *manufacturing* in *wholesale* and *retail* trade and in transport, storage and communication. The five sector considered account for approximately 90% of private sector employment. Results are available on request.



Private sector, worker's wage in logs. Non-parametric estimates: normal kernel, rule of thumb. Results displayed on the 1-99 percentile range.

Fig. 3 Wage distribution by highest completed education, Nov. 2000

Figure 3 displays non parametric estimates of the (log) wage distribution for workers that are in the panel in the last week of November 2000. With the exception of primary school and high school workers, wage inequality is higher the higher the education level. The wage distribution for primary and high school workers is bimodal, possibly due to the minimum wage. Wage inequality for high school workers is significantly higher than for any other education category, that for primary school workers being the second highest. [28]

Figure 4 illustrates the (expected) wage profile, unconditional, and conditioning on the originating labor market state. The analysis is restricted to wages from private employment and the series is normalized by the sample mean in each year, to account both for inflation and productivity growth. Also, we trim the bottom and upper percentile of wage observations. Finally, consistent with the on-the-job search view, earnings received in future jobs are counted as part of the current job wage stream.

[28] The 90/10 percentile ratios for the education groups listed in Figure 3 are equal to 2.86, 3.54, 2.03, 2.07, 2.56 and 2.55, respectively.

[29]. Bobbio (2010) checks the robustness of the statistics discussed below along several dimensions and complement the picture by looking at within-job wage changes.

The unconditional wage profile is steeper the higher the education level, with the exception of high school and primary school workers, who experience the highest and second highest relative increase in wages over time, respectively.[30] The ranking across groups is the same as for wage inequality, which is consistent with the on-the-job search explanation of wage dispersion – abstracting from differences across groups in the underlying distribution of job, ,*firm* and other worker characteristics.

However, most of the increase in wages over time is concentrated in the first 10 years of working life and, according to estimates by Bagger et al. (2014), can be attributed to the accumulation of human capital. We compute the profile separately for workers age 15-30 and age 35-45 (at job onset). Pooling together observations for workers with different education levels, the relative wage increase after 10 years of continued employment equals 42.2% in the first case and 4.1% in the second case. Across education groups, the unconditional profile for primary and high school workers age 35-50 is substantially flatter than for workers 15-30, the percentage increase being 0.0 and 13.1% respectively, instead of 55.4 and 45.2%.[31,32] The worker may have reached the top of the "wage ladder" by the time s/he is 35, which is consistent with the wage growing at the same rate as productivity (a flat profile). However, the profile is flat for workers going through unemployment as well. Indeed, for workers 35-45 the wage profile following unemployment is flatter than the unconditional profile (except than for those with a postgraduate education) suggesting that heterogeneity and unemployment scaring may play an important role – Ruhm (1991).

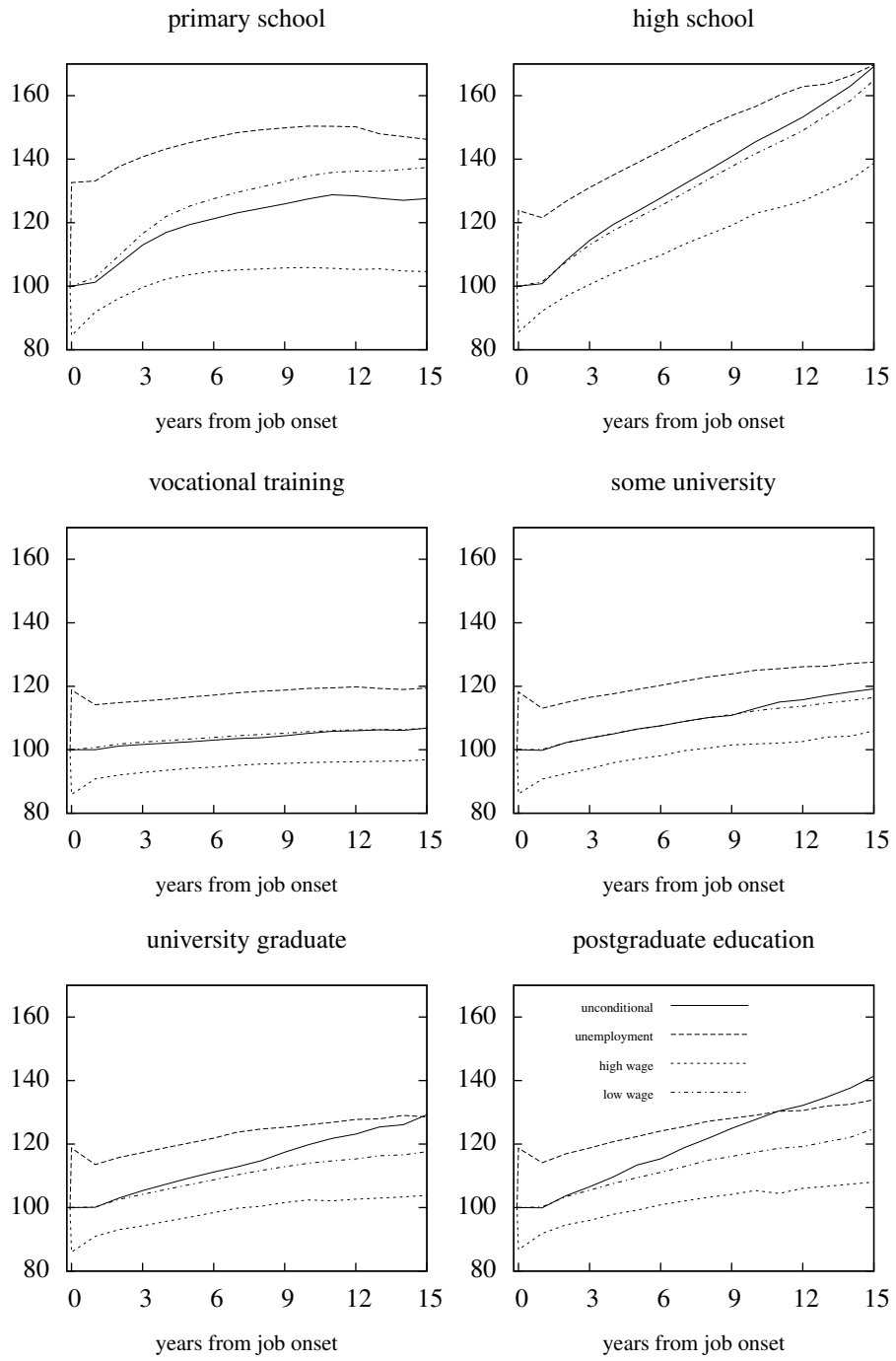
Also, wage cuts between jobs are large and pervasive: looking at the whole sample, the probability of a wage cut is 41.3% and the wage decreases by 14.9% on

[29] The normalization is made relative to the mean wage for the whole private sector, i.e. before splitting the population by age segment, and before trimming. For each job we record the wage in the first year and then in each year that follows, until the worker enters unemployment or exits the labor market. Also, if available, we record the previous state: unemployment, employment at a lower or higher wage. Then, we compute the profile relative to the first wage observation and average across profiles. Suppose a worker is hired from unemployment in period 0, changes job at 2 experiencing a wage increase and then becomes unemployed in period 6; the wage is observed in each of the six periods, from 0 to 5. The employment spell counts as two profiles, the first contributing to unemployment bins 1 to 5, $(\frac{w_1}{w_0},...,\frac{w_5}{w_0})$, and the second contributing to "employment +" bins 1 to 3 $(\frac{w_3}{w_2},...,\frac{w_5}{w_2})$.

[30] After 10 years of continued employment, the relative wage increases for the education groups listed in Figure 4 are 34.7, 41.7, 5.7, 12.3, 14.0 and 17.5%, respectively.

[31] The percentage increases for workers age 35-45 in the remaining four education groups listed in Figure 4 are 3.8, 11.2, 12.4 and 15.3, respectively. Detailed figures by originating labor market state, age and highest completed education are available on request.

[32] For the U.S., Altonji and Shakotko (1987) find small to no returns to tenure, while Topel (1991) estimates are in excess of 20% over 10 years. Reassessing the debate Altonji and Williams (2005) settle on a figure of 10%. Descriptive work by Rubinstein and Weiss (2006) on U.S. data shows that, after 10 years in the labor market, the wage increases by approximately 50% and 80% for high school and university graduates respectively.



Wage profile expected at job onset (time 0); unconditional and conditioning on the originating labor market state: unemployment, employment at a higher or lower wage. The wage series is normalized by the mean in each year. The wage at onset, or in the previous job, is set to 100. Wages in future jobs enter the current job wage stream – for details see footnote 29. Highest completed education as recorded at job onset.

Fig. 4 Wage profile by originating labor market state and education

average. [33, 34] Following a cut, the wage remains persistently low and the gap with the unconditional profile does not narrow over time. Thus, such job changes appear to be motivated by reasons other than the prospect of future higher wages – Postel-Vinay and Robin (2002). Taber and Vejlin (2011) advance the idea that compensating differentials may be important to explain these facts, as well as wage dispersion. Bobbio (2010) extends the on-the-job search framework to allow for replacement search on the side of *firms* replacement search may counter on-the-job search to generate high mobility and wage dispersion with low overall wage increases; also, low ability workers may prefer safer, but lower paying jobs.

6 Flows

Figure 5 displays life table estimates of the unconditional unemployment and job hazard (for the private sector) by highest completed education. The data is grouped at the weekly and quarterly level respectively and for each education group estimates are decomposed by ending state under the competing-risk assumption.[35]

The unconditional hazard displays strong negative duration dependence for all education groups. In addition, it displays periodicity, at the monthly level and yearly level. This may be due to misreporting by *employers* regarding the beginning and end of the employment relation. Alternatively, it may reflect genuine periodicity in the timing of employment contracts. Even in this case, the periodicity suggests that the starting date at the new job may be a poor proxy for the failure time of the underlying search process. Finally, with regard to the job hazard, transitions to another job are more likely than transitions to unemployment.

With the exception of high school graduates, workers with a higher education level find a job more easily and remain on the job longer; conditional on experiencing a separation, they are more likely to move to another job without transitioning to unemployment. Overall, the patterns for wages and flows across education groups suggests that workers with a higher education level face a higher offer arrival rate and more homogeneous job offers and, while on the job, exploit the between *employer* competition to obtain higher wages. High school workers display a substantially higher rate of labor market turnover, both in and out of unemployment and across jobs. From an on-the-job search perspective, a high job-to-job transition rate is coherent with a steep wage profile and high wage dispersion.[36]

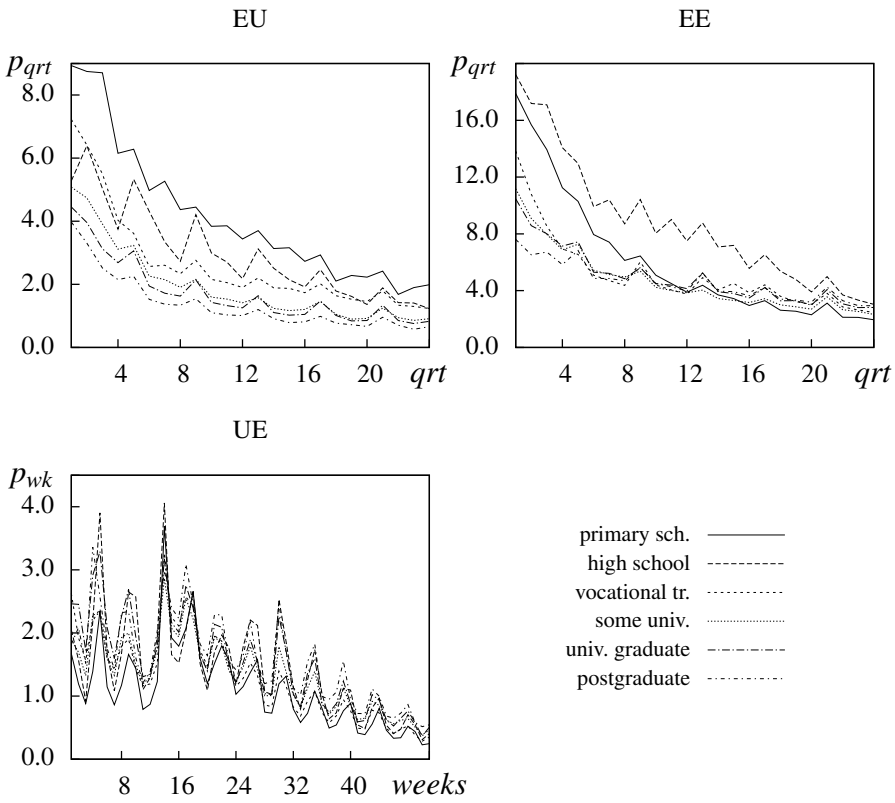
---

[33] The probability of a wage cut between jobs is 39.4, 36.4, 45.8, 43.7, 42.4 and 40.0% for each education group listed in Figure 4 respectively.

[34] Postel-Vinay and Robin (2002) and De Melo (2012) document similar figures using French and U.S. data respectively.

[35] Spells ending with a transitions to self-employment, non-participation or retirement are treated as censored spells.

[36] For each education group listed in Figure 5, mean job duration is 88, 75, 151, 185, 182 and 214 weeks, mean unemployment durations is 136, 81, 100, 101, 80, and 86 weeks and, finally, the



Life table estimates of the quarterly job hazard (top) and weekly unemployment hazard (bottom) decomposed by ending state under the competing risk assumption.

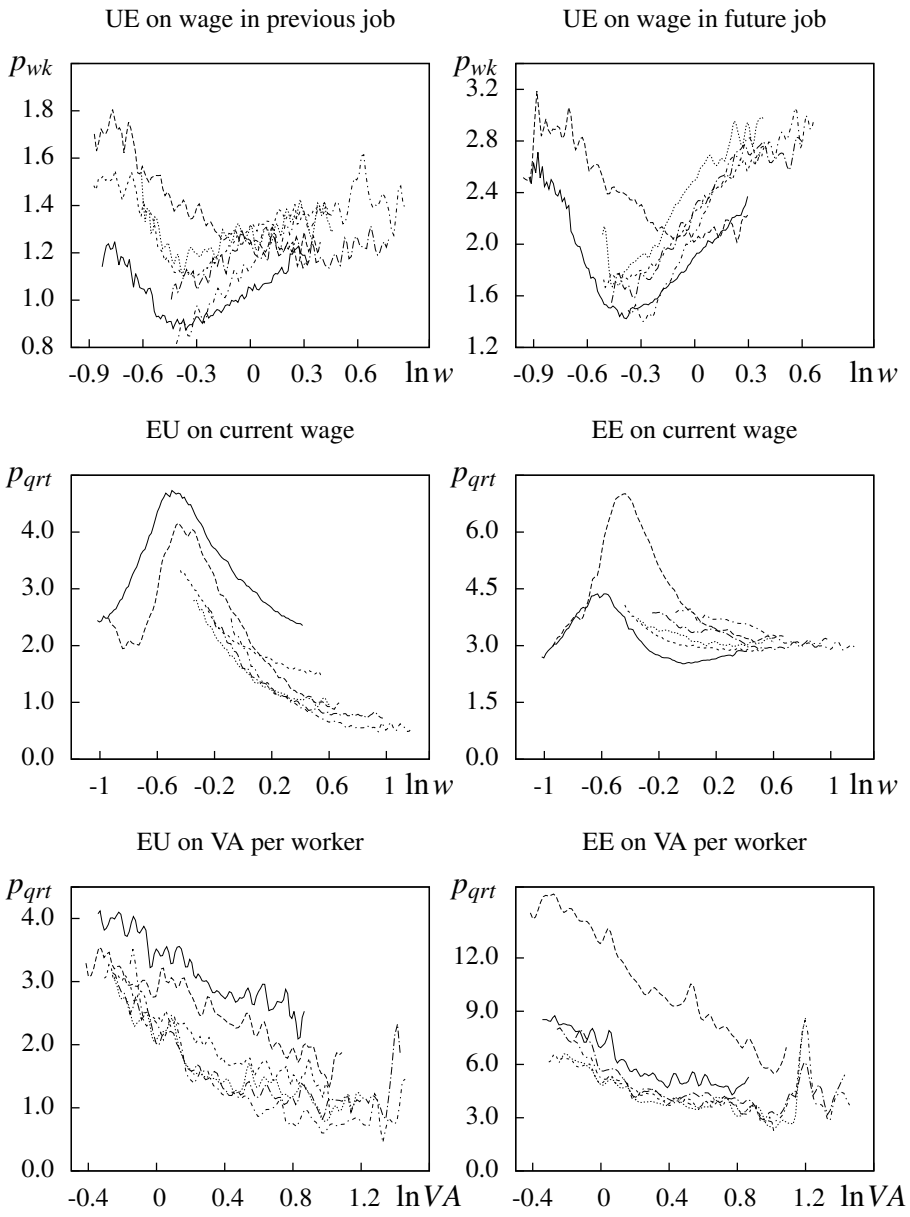
Fig. 5 Job and unemployment hazard, Kaplan Meier estimates

Figure 6 displays results for the local linear regression of a dummy on spell covariates, for the same education groups. For job spells the regressor is the worker's wage, or the *employer*'s value added per FTE ; for unemployment spells it is the wage in the previous or following job.[37]

Except than for workers earning low wages and, particularly, primary and high school workers, the relation between the wage and the probability of becoming unemployed or that of moving to another job are both negative; higher past and future

fraction of separations explained by job-to-job transitions (under the competing-risk assumption) is 0.41, 0.56, 0.56, 0.63, 0.66 and 0.72.

[37] Nominal *quantities* are mean-detrended. For jobs ending between January and November we report the wage recorded in the previous year (the wage is observed in the last week of November of each year). We use a normal kernel and the rule-of-thumb for the bandwidth choice. Results are displayed for the $5-95$ percentile range of the regressor.



Local linear regression of the weekly unemployment hazard (first row) and of the quarterly job hazard on employee's wage (second row) and employer's value added per FTE (third row) decomposed by ending state under the competing risk assumption. Nominal quantities are mean-detrended. Normal kernel and rule-of-thumb, results displayed on the 5-95 percentile range.

Fig. 6 Job and unemployment hazard, wages and labor productivity

wages are positively correlated with the probability of finding a job.[38] More productive *firms* and firms paying higher wages have better retention rates, due to both a lower transition rate to other jobs and to unemployment.[39]

The on-the-job search model with homogeneous workers naturally generates a negative correlation between the job-to-job transition rate and the worker's wage, or value added per worker, if *firms* are heterogeneous. Extending the framework to allow for idiosyncratic productivity shocks reinforces the correlation, as worker may anticipate and escape unemployment by moving to another *firm* – Nagypál (2005). Such model also generates a negative correlation between the rate of transitions to unemployment and the worker's wage, or value added per worker – Mortensen and Pissarides (1994). The relation may be explained by learning and selection as well – Jovanovic (1979). Allowing for worker heterogeneity and endogenous search intensity, the model generates a positive correlation between the wage and the unemployment hazard, as better workers face stronger incentives to search – Christensen et al. (2005). However the relation may also be explained by human capital depreciation and unemployment scaring – Ruhm (1991) –, or by firm heterogeneity if jobs are scarce, so that *firms* select better workers out of unemployment – Bobbio (2010) and Lise et al. (2013).

As noted above the job and unemployment hazards do not very monotonically with the wage. The on-the-job search framework can be reconciled with these features of the data if workers are heterogeneous: *conditional* on a worker type, the job-to-job transition rate declines with the worker's wage; however, if better workers earn higher wages and change job more frequently, then the *unconditional* hazard of moving to another job can increase and then decline with the wage.

Finally, we turn to the analysis of flows over the business cycle, which has received much attention in reaction to Shimer (2005)'s work. The interested reader is referred to Jonas Staghoej contribution in this volume for a more exhaustive treatment of the topic. Shimer notes that the textbook search and matching model has difficulties replicating the high volatility of market tightness (the vacancy-to-unemployment ratio). When unemployment rises during a recession, workers' outside options rapidly deteriorate, which should raise profits and prompt the creation of new vacancies. As a result, the *ue* rate should be mildly countercyclical. Instead, during recessions the unemployment pool increases without a corresponding increase in the flow out of unemployment and into employment. [40]

---

[38] The schedules in the top-right panel are higher than the corresponding schedules in the top-left panel, because the wage in the future job is not observed if the unemployment spell is censored, and censored spell tend to be longer. Similarly, the job hazards in Figure 6 are lower than in Figure 5 because no wage is observed for job spells lasting less than year and ending before November.

[39] When using the firm average wage as the regressor we obtain similar results as for value added per FTE . Also, we stratify the analysis by sector: the job hazard of transitioning to another job or to unemployment both decline with firm average wage or value added per FTE , for all industries considered in footnote 27, except *construction*. Results are available on request.

[40] Since then, Hagedorn and Manovskii (2008) have confuted this claim: if the utility associated with being unemployed is sufficiently high, a worker's outside option is inelastic to aggregate changes in productivity.

The observation suggests that understanding job creation is important to explain the cyclical behavior of unemployment, while research has emphasized the role of job destruction – Darby et al. (1985), Blanchard and Diamond (1990) and Davis et al. (1996). Subsequent empirical work has tried to assess the relative importance of the "ins and outs" of unemployment as determinants of unemployment volatility – Shimer (2007), Hall (2005), Fujita and Ramey (2009) and Elsby et al. (2009). Two aspects have received particular attention: the decomposition of the unemployment variance; and the cyclicality of the transition rate from employment to unemployment or, alternatively (and not equivalently), of the separation rate. The discussion has focused on measurement issues – e.g. time aggregation, sample attrition, misreporting, etc. – and the appropriate definition of the relevant rates; the problem arises when departing from the simple bivariate case to account for job-to-job transitions and non-participation.

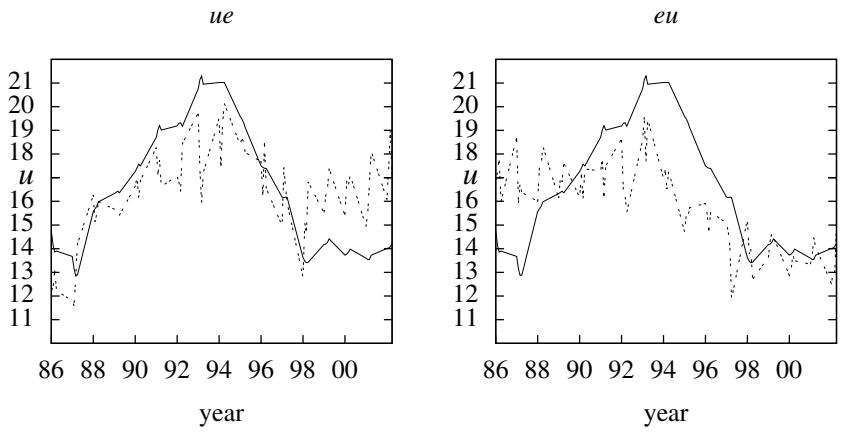
The literature has (mainly) used CPS data for the U.S. The Danish data offers some advantages: it has weekly frequency and covers the whole population, which mitigates some of the measurement issues mentioned above. Second, it provides detailed information on a worker's labor market state. However, as noted above, it should be kept in mind that the distinction between non-participation and unemployment, or that between lay-offs and job-to-job transitions is problematic, especially when a worker is not officially registered as unemployed. We revisit Shimer (2007) and Fujita and Ramey (2009) work using the Danish data. We restrict attention to the "gross flow" approach to compute transition rates.

For each week between 1985 and the end of 2003 we compute the pool of employed, unemployed and inactive workers (retired or non-participating), as well as the six gross flows from one state to another. Next, series are aggregated at the monthly level, the first and last year are truncated, and the remaining data is seasonally adjusted using a ratio to moving average technique.[41] Finally, we compute transition rates by dividing the number of transitions by the pool size of the state of departure. Shimer (2007) finds that the steady state condition for unemployment in the bivariate case (where the worker can be only employed or unemployed) closely tracks actual unemployment:

$$u_t \approx \frac{eu_t}{eu_t + ue_t} \tag{1}$$

In the Danish data the correlation between the two series is .90, which is lower than the number found by Shimer (2007) for the CPS data. Figure 7 displays quarterly averages of the unemployment series (solid) against the right hand side of (1), keeping constant at its average level the *eu* rate (left panel) or the *ue* rate (right panel). The left panel is interpreted as the contribution of the *ue* rate to unemployment volatility, $u_t^{ue} \equiv \bar{eu}/(\bar{eu} + ue_t)$; similarly for the right panel. Next, we detrend the three series using the HP filter with penalty parameter equal to 1600 and, as in Shimer (2007), we quantify the contribution of the *ue* rate by computing the ratio $cov(u_t, u_t^{ue})/var(u_t) = .625$; similarly, for the *eu* rate we com-

[41] The first year is removed because persons are gradually included in the spell data in the first few weeks of 1985. The last year is removed because the coding of the employment state generates a spurious increase in unemployment, as discussed in Section 4.2.



Actual unemployment series (solid) against the right hand side of (1) (dashed). In the left (right) panel, the *eu* (*ue*) rate is kept constant at its average level. Data is aggregated at the monthly level and seasonally adjusted; then, rates are computed by dividing the gross flow by the corresponding pool size and quarterly averaged.

Fig. 7 Ins and outs, contribution to unemployment volatility

pute $cov(u_t, u_t^{eu})/var(u_t) = .313$.[42] Together the two numbers sum to .938. Also, following Shimer (2007), we consider the trivariate system accounting for non-participation. The figure (not reported) is unaltered, because the other four rates are essentially acyclical. The decomposition, delivers .543 and .281 respectively.[43]

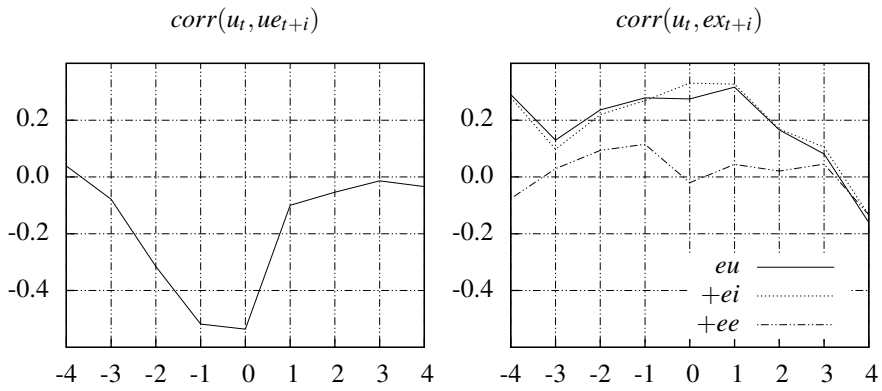
Next, as in Fujita and Ramey (2009) we compute the simple correlation at different lags and leads between the actual unemployment series u_t and the ue_t and eu_t series – left and right panel of Figure 8 respectively; the three series are quarterly averaged and then HP detrended. Regarding the separation rate, we also include statistics for the $eu_t + ei_t$ and the $eu_t + ei_t + ee_t$ rates – for comparability with Hall (2005) and Elsby et al. (2009).

Overall, findings are remarkably close to those in Shimer (2007), Hall (2005), Fujita and Ramey (2009) and Elsby et al. (2009) using CPS data. The variance decomposition attributes two thirds of the unemployment variance to movements in the *ue* rate; the *ue* rate is procyclical, while the *eu* rate is countercyclical; the *ee* rate is procyclical, so the overall separation rate is almost acyclical. Transitions in and out of non-participation are quite acyclical as well, so accounting for non-participation hardly changes the results.

While the analysis suggests that both the entry and exit rates into and out of unemployment are important to explain its cyclical behavior, the results do not re-

[42] Fujita and Ramey (2009) and Elsby et al. (2009) devise different decompositions, which however share the same spirit and lead to similar results.

[43] The contribution of the other four rates ranges from .009 to .046 and the six numbers sum to .939.



Correlation at different lags and leads between the unemployment rate and the *ue* rate (left) and the three components of the separation rate (to unemployment, to inactivity or to another job) added together one by one. The data is aggregated at the monthly level, and then seasonally adjusted. Rates are computed as the gross flow divided by the corresponding pool size. Series are then quarterly averaged and HP detrended ($\lambda = 1600$).

Fig. 8 Ins and outs, business cycle comovement

solve the debate on their relative importance. This is essentially because the variance decomposition is reduced form. The assessment should be based on the counter factual: what would unemployment look like if we could intervene to alter one of the two factors? Shimer (2007)'s approach seems to do that, by keeping one rate constant and letting the other follow the historical series. But, in practice, the *ue* rate declines sharply in recessions because the unemployment pool (the denominator) increases, and not because the gross flow out of unemployment (the numerator) drops – Kennan (2005). This is true for the Danish data as well. Nonetheless, Shimer (2007)'s point remains valid: even if lays-offs are abnormally high at the onset of a recession, the reference model of the labor market, does not explain why workers who become unemployed are not quickly re-absorbed into employment.

7 Conclusions

The Danish data constitutes a unique source of information due to its quality and its extent. The data has the potential to become a laboratory for social studies, but access has been limited by the lack of systematic documentation. This chapter provides an introduction and establishes basic statistics that can serve as a reference. In addition, it explores in more detail wages and worker flows, for which the data is uniquely well suited.

Wage inequality is low, but dispersion in productivity is large. Higher wages are associated with larger and more productive *employers*, but the relation between productivity and wage is concave. Regarding wage dynamics, seniority effects are mild or nil, for workers age 35 and older. Wages move substantially between jobs, but drops are large and frequent, and the wage remains substantially lower for an extended period of time. Seniority effects are even lower following unemployment. Wage inequality and seniority effects both increase with education, except than for primary and high school workers.

The job and unemployment hazards display strong negative duration dependence and are negatively and positively correlated with the wage, respectively, except than for low wage workers.

The unconditional job and unemployment hazard display strong negative duration dependence. The wage is negatively correlated with the probability of becoming unemployed or of transitioning to another job and higher wages are associated with shorter unemployment duration, except than for low wage workers and particularly primary and high school workers. The Danish data supports findings for the U.S. in the debate on the cyclical behavior of the rates in and out of unemployment. The *ue* rate explains 2/3 of unemployment volatility – though one should be careful not to attach any particular structural interpretation to this number. The *ue* rate is strongly procyclical, while the *eu* rate is strongly countercyclical. The *ee* rate is procyclical so, overall, the separation rate is essentially acyclical.

References

Altonji J, Shakotko R (1987) Do wages rise with job seniority? The Review of Economic Studies 54(3):437–459

Altonji J G, Williams N (2005) Do wages rise with job seniority? a reassessment. Industrial and Labor Relations Review 58(3):370–397

Bagger J, Bunzel H, Christensen B J (2016) On the determinants of wage and productivity dispersion. In: Bunzel H, Christensen B J, Mortensen D T, Sørensen K L (eds.), Firm Heterogeneity, Labor Markets and International Trade: Evidence from Danish Matched Employer-Employee Data. Heidelberg: Springer

Bagger J, Christensen B, Mortensen D (2013) Wage and labor productivity dispersion: The roles of total factor productivity, labor quality, capital intensity, and rent sharing. Tech. rep., Working Paper

Bagger J, Fontaine F, Postel-Vinay F, Robin J M (2014) Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics. American Economic Review 104(6):1551–1596

Bagger J, Lentz R (2014) An empirical model of wage dispersion with sorting. NBER Working Paper No. 20031

Bertheau A (2016) Unpublished Tech. note, Aarhus University

Bjelland M, Fallick B, Haltiwanger J, McEntarfer E (2011) Employer-to-employer flows in the united states: Estimates using linked employer-employee data. Journal of Business & Economic Statistics 29(4):493–505

Blanchard O, Diamond P (1990) The cyclical behavior of the gross flows of US workers. Brookings Papers on Economic Activity 1990(2):85–155

Bobbio E (2010) Replacement search and wages. Bank of Italy

Bunzel H (2016) Econau data sets. In: Bunzel H, Christensen B J, Mortensen D T, Sørensen K L (eds.), Firm Heterogeneity, Labor Markets and International Trade: Evidence from Danish Matched Employer-Employee Data. Heidelberg: Springer

Bunzel H, Hejlesen M (2016) Unpublished Tech. note, Aarhus University

Burdett K, Mortensen D (1998) Wage differentials, employer size, and unemployment. International Economic Review 39(2):257–273

Christensen B, Lentz R, Mortensen D, Neumann G, Werwatz A (2005) On-the-job search and the wage distribution. Journal of Labor Economics 23:31–58

Darby M, Haltiwanger J, Plant M (1985) Unemployment rate dynamics and persistent unemployment under rational expectations. The American Economic Review 75(4):614–637

Davis S, Haltiwanger J, Schuh S (1996) Job creation and destruction. MIT press

De Melo R (2012) The Implications of Search Models for Wage Dynamics: an Empirical Assessment. University of Chicago mimeo

Elsby M, Michaels R, Solon G (2009) The ins and outs of cyclical unemployment. American Economic Journal: Macroeconomics 1(1):84–110

Fujita S, Ramey G (2009) The cyclicality of separation and job finding rates. International Economic Review 50(2):415–430

Hagedorn M, Manovskii I (2008) The cyclical behavior of equilibrium unemployment and vacancies revisited. The American Economic Review 98(4):1692–1706

Hall R (2005) Job loss, job finding, and unemployment in the US economy over the past fifty years. NBER Macroeconomics Annual 101–137

Jovanovic B (1979) Job matching and the theory of turnover. The Journal of Political Economy 87(5):972–990

Kennan J (2005) Job Loss, Job Finding, and Unemployment in the US Economy over the Past Fifty Years: Comment. NBER Macroeconomics Annual 159–164

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

Lise J, Meghir C, Robin J (2013) Mismatch, sorting and wage dynamics. NBER Working Paper Series 18719

Mortensen D (2003) Wage dispersion: why are similar workers paid differently? Mit Press

Mortensen D, Pissarides C (1994) Job creation and job destruction in the theory of unemployment. The Review of Economic Studies 61(3):397–415

Nagypál É (2005) On the extent of job-to-job transitions. Northwestern University mimeo

Postel-Vinay F, Robin J (2002) Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica 70(6):2295–2350

Rosholm M, Svarer M (2004) Endogenous wage dispersion in a search-matching model. Labour Economics 11(5):623–645

Rubinstein Y, Weiss Y (2006) Post schooling wage growth: Investment, search and learning. Handbook of the Economics of Education 1:1–67

Ruhm C (1991) Are workers permanently scarred by job displacements? The American Economic Review 81(1):319–324

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

——— (2007) Reassessing the ins and outs of unemployment. NBER Working Paper Series 13421

Taber C, Vejlin R (2011) Estimation of a Roy/Search/Compensating Differential Model of the Labor Market. Aarhus University mimeo

Topel R (1991) Specific capital, mobility, and wages: Wages rise with job seniority. Journal of Political Economy 99(1):145–176

Economics Working Papers

2017-04: Michael Koch and Marcel Smolka: Foreign Ownership and Skill-biased Technological Change

2017-05: Mette Trier Damgaard and Helena Skyt Nielsen: Nudging in education: A survey

2017-06: Alexander K. Koch and Julia Nafziger: Motivational Goal Bracketing: An Experiment

2017-07: Ina C. Jäkel Allan Sørensen: Exporter Price Premia?

2017-08: Marianne Simonsen, Lars Skipper and Niels Skipper: Piling Pills? Forward-Looking Behavior and Stockpiling of Prescription Drugs

2017-09: Federico Ciliberto and Ina C. Jäkel: Superstar Exporters: An Empirical Investigation of Strategic Interactions in Danish Export Markets

2017-10: Anna Piil Damm, Britt Østergaard Larsen, Helena Skyt Nielsen and Marianne Simonsen: Lowering the minimum age of criminal responsibility: Consequences for juvenile crime and education

2017-11: Erik Strøjer Madsen: Branding and Performance in the Global Beer Market

2017-12: Yao Amber Li, Valerie Smeets and Frederic Warzynski: Processing Trade, Productivity and Prices: Evidence from a Chinese Production Survey

2017-13: Jesper Bagger, Espen R. Moen and Rune M. Vejlin: Optimal Taxation with On-the-Job Search

2018-01: Eva Rye Johansen, Helena Skyt Nielsen and Mette Verner: Long-term Consequences of Early Parenthood

2018-02: Ritwik Banerjee, Nabanita Datta Gupta and Marie Claire Villeval: Self Confidence Spillovers and Motivated Beliefs

2018-03: Emmanuele Bobbio and Henning Bunzel: The Danish Matched Employer-Employee Data