



# THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

# Essays on Long-term Unemployment in Spain

*Cristina Lafuente Martinez*

Doctor of Philosophy



The University of Edinburgh  
School of Economics

2017

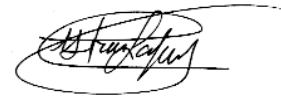


# Declaration of Own Work

I declare that this thesis was written and composed by myself and is the result of my own work unless clearly stated and referenced. This thesis has not been submitted for any other degrees or professional qualifications.

The work presented in Chapter 2 is based on my work with my supervisor, Maia Güell. Maia has agreed that the essay can appear within this thesis, and that it represents a substantial contribution on my part. In particular, I carried out the data handling, main statistical work, adapted the methodology in section 2.2, provided the description in section 2.3.5, contributed to the description of the data in section 2.3 and to the analysis in section 2.4.

Cristina Lafuente Martinez



Date: 31/12/2017

*to Jim, my parents, Ana and Cecilia Elisa*

*Without you this thesis would not have been possible*

# Acknowledgments

I am very grateful to my supervisors Maia Guell and Ludo Visschers for their guidance throughout these years. I have learnt a great deal from both of them. I would also like to thank professor Mike Elsby for his supervision on my first year and Philip Kircher for his support and wise comments. Finally, I would also want to thank professor Rios-Rull for kindly inviting me to the University of Pennsylvania for my third year and Iouri Manovskii and Ken Burdett for useful comments and their kind support while at Penn.

I also wish to thank my fellow PhD students for their unconditional support and helpful comments. In particular, I would like to thank Yulia Moiseeva, Sergei Plekhanov, Di Yang, Julia Faltermeier, Carl Singleton, Rachel Forshaw, Aspasia Bisopoulou, Tomasz Sulka and the second year's of Penn. Furthermore, I am grateful to the PhD directors and administrative staff for making my PhD journey as smooth as possible. My work has been financially supported by the ESRC [grant number ES/J500136/1].

Special thanks to Adam Granduciel for the mental and moral support, especially through the final stages of the thesis.



# Abstract of Thesis

This thesis is comprised of three essays relating to long term unemployment in Spain. The first chapter is methodological analysis of the main dataset that is used throughout the thesis. The second and third chapter provide two applications of the dataset for the study of long term unemployment. The methodology in these chapters can be easily adapted to study unemployment in other countries.

*Chapter 1. On the use of administrative data for the study of unemployment* Social security administrative data are increasingly becoming available in many countries. These are very attractive data as they have a long panel structure (large N, large T) and allow to measure many different variables with higher precision. Because of their nature they can capture aspects that are usually hidden due to design or timing of survey data. However, administrative data are not ready to be used for labour market research, especially studies involving unemployment. The main reason is that administrative data only capture those registered unemployed, and in some cases only those receiving unemployment benefits. The gap between total unemployment and registered unemployment is not constant neither across workers characteristics nor time. In this paper I augment Spanish Social Security administrative data by adding missing unemployment spells using information from the institutional framework. I compare the resulting unemployment rate to that of the Labour Force Survey, showing that both are comparable and thus the administrative dataset is useful for labour market research. I also explore how the



administrative data can be used to study some important aspects of the labour market that the Labour Force survey can't capture. Administrative data can also be used to overcome some of the problems of the Labour Force survey such as changes in the structure of the survey. This paper aims to provide a comprehensive guide on how to adapt administrative datasets to make them useful for studying unemployment.

*Chapter 2. Unemployment Duration Variance Decomposition à la ABS: Evidence from Spain* Existing studies of unemployment duration typically use self-reported information from labour force surveys. We revisit this question using precise information on spells from administrative data. We follow the recent method proposed by Alvarez, Borovickova and Shimer (2015) for estimating the different components of the duration of unemployment using administrative data and have applied it to Austria. In this paper we apply the same method (the ABS method hereafter) to Spain using Spanish Social Security data. Administrative data have many advantages compared to Labour Force Survey data, but we note that there are some incompleteness that need to be enhanced in order to use the data for unemployment analysis (e.g., unemployed workers that run out of unemployment insurance have no labour market status in the data). The degree and nature of such incompleteness is country-specific and are particularly important in Spain. Following Chapter 1, we deal with these data issues in a systematic way by using information from the Spanish LFS data as well as institutional information. We hope that our approach will provide a useful way to apply the ABS method in other countries. Our findings are: (i) the unemployment decomposition is quite similar in Austria and Spain, specially when minimizing the effect of fixed-term contracts in Spain. (ii) the constant component is the most important one; while (total) heterogeneity and duration dependence are roughly comparable. (iii) also, we do not find big differences in the contribution of the different components along the business cycle.

*Chapter 3. Search Capital and Unemployment Duration* I propose a novel mechanism called search capital to explain long term unemployment patterns across different ages: workers who have been successful in finding jobs in the recent past become more efficient at finding jobs in the present. Search ability increases with search experience and depreciates with tenure if workers do not search often enough. This leaves young (who have not gained enough search experience) and older workers in a disadvantaged position, making them more likely to suffer long term unemployment. I focus on the case of Spain, as its dual labour market structure favours the identification of search capital. I provide empirical evidence that search capital affects unemployment duration and wages at the individual level. Then I propose a search model with search capital and calibrate it using Spanish administrative data. The addition of search capital helps the model match the dynamics of unemployment and job finding rates in the data, especially for younger workers.



# Lay Summary

Long Term Unemployment (unemployment of more than 12 months of duration, LTU thereafter) is a very serious issue. It negatively affects not only job market outcomes in the future but also health and well-being of workers who suffer it. The last recession has seen LTU rise in many countries but it has especially affected Southern European countries. Spain is one of them, and it's an interesting case in its own because of the large reforms that have taken place in the last 30 years, changing the structure of the labour market. In particular, the introduction of temporary contracts and its widespread use made Spain an anomaly among European countries with large labour market flows and highly protected employment. This increase in flows out of unemployment contributed towards the fall of long-term unemployment in the 1993-2007 period, from over 50% to below 20% in 2007. The financial crisis and the European debt crisis quickly reversed these patterns, with long-term unemployment reaching over 50% again in 2014. This increase was long-lived as well, as of 2017 it is still over 40%. One new and worrying feature of this increase is that it has disproportionately affected the young. Most explanations in the literature for the rise and persistence of LTU during recessions are targeted towards older workers – mainly loss of human capital and search incentives. In this thesis I turn to the data to further examine this issue.

My first chapter is devoted to the data themselves. In particular, I focus on the issues arising from using administrative datasets, which have complete working his-

tories of a large sample of workers, for the study of unemployment. Researchers have traditionally turned to Labour Force Surveys, as it contains self-reported information on whether workers are unemployed, but they have their shortcomings. In particular, attrition (non-response of individuals after the first interview) is an important issue for measuring unemployment in the Spanish Labour Force Survey. I show how the administrative and survey datasets are comparable and thus how administrative datasets can be used for the study of unemployment. This is crucial for long-term unemployment, since the information on duration of the administrative data is much more precise. Chapters 2 and 3 use these data to explore two questions that relate to long-term unemployment: how much changes in composition (heterogeneity) and duration dependence contribute towards the explaining the variance of the duration of unemployment, and a particular channel through which previous work experiences can affect unemployment duration.

The second chapter (co-authored with my supervisor Maia Güell) uses the expanded administrative dataset to decompose the variance of unemployment duration into three fundamental components: worker heterogeneity (observed and unobserved), duration dependence and constant hazard component. Worker heterogeneity refers to all individual characteristics that make workers experience different unemployment durations from others, like ability, age or gender. Duration dependence relates to the empirical observation that the odds of leaving unemployment are decreasing as individual's time in unemployment increases. It can be due to discrimination on the side of the employer, demoralising or other driving forces. Finally, the constant hazard component is the underlying probability that any unemployed worker, independent of individual characteristics or time in unemployment, finds a job. We interpret this as the overall (or macroeconomic) state of the labour market that affects all workers. We carry out the decomposition of unemployment variance in these three components following the methodology in Alvarez, Borovickova and Shimer (2015). This methodology requires a large sample with only two completed

spells of unemployment per individual. The administrative data from Chapter 1 is well suited for this. We find the constant hazard component to be the largest, explaining 60% of the total variation, with the rest split evenly between heterogeneity and duration dependence. We also find that the weight of the three components for Spain is remarkably similar to the one Alvarez, Borovickova and Shimer found for Austria, a European country with very different unemployment dynamics. Moreover, using the last recession we find that the weights of the different components do not substantially change over the business cycle. This result is very important for the study of long-term unemployment, as duration dependence and heterogeneity have both been proposed to explain its nature. Our results suggest that most of its volatility is due to aggregate fluctuations in the job finding rate, and that Spain has no fundamental difference in these components to Austria.

The third chapter seeks to explain part of the unobserved heterogeneity driving the increase of long-term unemployment during recessions. I investigate how past employment experiences can affect unemployment outcomes in the present. I propose a new channel, *search capital*, which comprises all the skills that help with job search and that are learnt through participating in the labour market as a job seeker. Thus workers who are more up-to-date with the job market are better at finding jobs than first job seekers or those who have enter unemployment from very long, stable jobs. This can lead to increasing differences in unemployment durations if both types of worker are in direct competition for few jobs, as it happens in large recessions where long tenured workers lose their jobs. I provide some evidence of the effect of search capital differences on the duration of unemployment spells. I do so by exploiting the variation of people who experience unemployment often (those in temporary contracts) and those who don't (in permanent contracts) in Spain to identify empirically search capital. I find that the more jobs a worker had in the past, the shorter her unemployment spells are. I also find a positive correlation between number of jobs and future wages. While these results speak of individual effects, I

build a dynamic search model to assess the aggregate effects of search capital. The model incorporates elements of two different literatures: risk aversion and savings from the *'turbulence'* literature and the dual labour market environment. This allows me to control for the two main competing theories that could also explain the correlations found in the empirical part: older individuals have more means to save assets and more human capital to lose and all individuals prefer a stable stream of income over a more volatile one. After calibrating the model using the Spanish administrative data, I find that search capital improves the fit of the model to the data, specially for younger workers. Search capital accentuates differences in unemployment duration across workers of different ages and helps explain why individuals in their late twenties-early thirties have the highest job finding rates of all - after which they steadily decline. Lucky workers who find stable jobs become worse searchers over time, but as long as they are employed this decrease in search capital is not reflected in the unemployment pool. There only the unlucky workers, who are still in temporary contracts, remain unemployed.

# Contents

<b>Declaration of Own Work</b>	<b>i</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>Abstract of Thesis</b>	<b>v</b>
<b>Lay Summary</b>	<b>ix</b>
<b>1 On the use of administrative data for the study of unemployment</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Data . . . . .	3
1.2.1 Description . . . . .	3
1.2.2 Prepare the MCVL for use . . . . .	8
1.3 Methodological Check . . . . .	12
1.3.1 The unemployment gap . . . . .	12
1.3.2 Closing the unemployment gap: LTU expansion . . . . .	15
1.3.3 Closing the unemployment gap: STU expansion . . . . .	17
1.4 Further Robustness checks . . . . .	23
1.5 Using the MCVL to enrich the LFS . . . . .	25
1.5.1 Attrition and Labour Market Flows . . . . .	26
1.5.2 Changes in Survey Design . . . . .	32
1.6 Conclusion . . . . .	33
Bibliography . . . . .	35



<b>2</b>	<b>Unemployment Duration Variance Decomposition à la ABS: Evidence from Spain</b>	<b>39</b>
2.1	Introduction . . . . .	39
2.2	The ABS method . . . . .	42
2.2.1	Estimation with two spells . . . . .	44
2.3	Data . . . . .	45
2.3.1	Data expansions . . . . .	46
2.3.2	Duration of unemployment in Administrative data . . . . .	46
2.3.3	In between employment spells . . . . .	50
2.3.4	On the spell number . . . . .	52
2.3.5	Austria and Spain . . . . .	55
2.4	Results . . . . .	57
2.4.1	Variance decomposition with raw data and LTU expansion . . . . .	59
2.4.2	Variance decomposition with all expansions . . . . .	60
2.4.3	Business cycle comparisons . . . . .	62
2.5	Conclusions . . . . .	64
	Bibliography . . . . .	66
<b>3</b>	<b>Search capital and Unemployment Duration</b>	<b>69</b>
3.1	Introduction . . . . .	69
3.2	Search Capital and Long Term Unemployment . . . . .	72
3.2.1	Search Capital, Dual Labour Markets and LTU . . . . .	73
3.2.2	Related explanations of LTU . . . . .	78
3.3	Empirical Analysis . . . . .	83
3.3.1	The Data . . . . .	83
3.3.2	Identification Strategy . . . . .	87
3.3.3	Empirical approach . . . . .	92
3.3.4	Results . . . . .	96
3.4	Theoretical Model . . . . .	109

3.4.1	Value functions . . . . .	113
3.4.2	Calibration . . . . .	117
3.4.3	Results . . . . .	123
3.5	Conclusion . . . . .	130
	Bibliography . . . . .	133
	Appendix A . . . . .	139
	Appendix B . . . . .	154
	Appendix C . . . . .	157



# List of Figures

1.1	Unemployment rate in Spain . . . . .	14
1.2	Unemployment rate by sex . . . . .	14
1.3	Unemployment rate by age . . . . .	14
1.4	Unemployment rates - LFS and LTU expanded Social Security . . . .	16
1.5	Age Distribution by Expansion . . . . .	19
1.6	Spells before and After Unemployment . . . . .	19
1.7	Unemployment rates - STU Expansion . . . . .	20
1.8	Unemployment rates by gender . . . . .	22
1.9	Unemployment rates by age group . . . . .	22
1.10	Relationship with the Employment Office . . . . .	24
1.11	Alternative unemployment rate series . . . . .	24
1.12	Attrition of Unemployment Stocks in the LFS . . . . .	26
1.13	Labour Market Flows in LFS . . . . .	28
1.14	Flows out of unemployment . . . . .	30
1.15	Flows into unemployment . . . . .	30
1.16	Quarterly Flows: Between contract types . . . . .	32
2.1	Duration in months of (uncompleted) unemployment in 2013 (III), LTU expansion . . . . .	48
2.2	Duration in weeks of completed unemployment spells, 2005-2013, LTU expansion . . . . .	49

2.3	Duration in months of (uncompleted) unemployment in 2013 (III), STU expansion . . . . .	53
2.4	Duration in weeks of completed unemployment spells, 2005-2013, STU expansion . . . . .	53
2.5	Unemployment and Long Term Unemployment, Austria and Spain . .	55
2.1	Hazard Rate By Number Of Spell . . . . .	58
2.2	Decomposition with LTU addition . . . . .	59
2.3	Decomposition with LTU and STU additions . . . . .	61
2.4	Spain along business cycle . . . . .	63
3.1	Youth and overall long term unemployment . . . . .	71
3.2	Youth and overall long term unemployment, annual log change . . . .	71
3.3	Long term unemployment in Spain, by age . . . . .	76
3.4	Flows into unemployment, by contract type . . . . .	76
3.5	Wage distributions . . . . .	91
3.6	Tenure-Wage distributions . . . . .	119
3.7	Monthly Transition Rates by Age (pre 2008, SS) . . . . .	121
3.8	Duration of unemployment by number of contracts and search capital level . . . . .	122
3.9	Search Capital by age . . . . .	124
3.10	Temporary share and Unemployment rate . . . . .	125
3.11	Transition rates . . . . .	126
3.12	Duration of unemployment, by search capital . . . . .	128
3.13	Transition rates without search capital . . . . .	129
A.1	Histogram of Spell duration, by Extension . . . . .	153

# List of Tables

1.1	Data Comparison Table . . . . .	7
1.1	STU Expansion spells, by type . . . . .	18
2.1	Duration of unemployment . . . . .	47
2.2	Number of unemployed . . . . .	50
2.3	Spell Number Correction . . . . .	54
2.4	Structure of Unemployment Insurance . . . . .	57
3.1	Descriptive Statistics . . . . .	86
3.2	Variables of interest . . . . .	87
3.3	Regressions on Unemployment Duration . . . . .	97
3.4	Probability of Long-Term Unemployment . . . . .	101
3.5	Regressions on Future Wages . . . . .	104
3.6	Regressions on duration of next job . . . . .	107
3.7	Calibration . . . . .	118
A.1	Summary Statistics of Unemployment Spells . . . . .	153
B.1	Variance decomposition with raw data and LTU expansion . . . . .	154
B.2	Variance decomposition with LTU and STU expansion . . . . .	155
B.3	Variance decomposition along the business cycle . . . . .	156
C.1	Regressions on Duration - controls . . . . .	158
C.2	Quadratic term for temporary contracts . . . . .	159
C.3	Regressions on Unemployment Duration, by next job industry . . . . .	160







# Chapter 1

## On the use of administrative data for the study of unemployment

### 1.1 Introduction

Administrative datasets are being increasingly used for labour market research (see for example Moffitt (1985), Katz and Meyer (1990), Sullivan and Von Wachter (2009), Tattara and Valentini (2010), Couch et al. (2011), Krueger and Mueller (2011), Bonhomme and Hospido (2017) among others). They offer many advantages over traditional Labour Force Surveys, from firm-worker identification to detailed and extensive working histories (large N, long T). However there are some challenges when using them to study unemployment. First, these data were not designed for research, but rather for administrative bookkeeping: who is making contributions to the social security and who is claiming benefits. This is especially true when unemployment benefits are contribution-based, so they are proportional to the social security contributions of the worker in her previous employment. Second, the definition of unemployment is not the same as the International Labour Office standard. Although the reception of benefits is in most cases conditioned on active search on the side of the worker, monitoring may not be perfect. Finally, in some countries the administration only keeps track of the unemployed while they are receiving benefits.

The aim of this chapter is to show how to implement two simple expansions in order to make administrative data ready to be used for the study of unemployment. First, unemployed workers who run out of benefits but keep searching will appear in the administrative data as if their spell was over. I add the missing days in between the end of a registered unemployment spell and the next employment spell to correct for this. Second, those who are not entitled to unemployment benefits because they had too short of a tenure in their previous job will also not appear as unemployed. Using the richness of the data and the institutional setting I identify these cases and add these spells into the data.

The main methodological check is to compare the resulting unemployment rate from the extended administrative data with that of the Labour Force Survey. To this aim I format the administrative dataset into a quarterly panel structure as in the Labour Force Survey. The results show evidence that the different expansions achieve their aim of adding the missing unemployment to the administrative dataset. In particular, the second expansion is crucial for youth unemployment and for women, who have a higher incidence of part-time and temporary contracts. As further check I use the information in the Labour Force Survey to reconstruct the administrative data's unemployment rate.

Finally I show how the expanded administrative data can help overcome some of the problems that the Labour Force Survey faces: attrition and changes in survey design. First, unemployed workers sometimes do not reply to two consecutive interviews, resulting in underestimated unemployment-to-unemployment flows. This affects exit rates from unemployment. Second, there was a major change in the survey in 2005 which did not affect stocks, although some flows were severely affected. The MCVL can help to distinguish which jumps correspond to actual events and which are artificially created as a result of the redesign of the survey.

The contribution of this chapter is twofold: First, it extends the methodology of García Pérez (2008) in adapting the administrative dataset in order to make it useful for research. I provide further systematic guidance using the institutional setting and the Labour Force Survey as a benchmark. Appendix A provides a detailed guide on how to do this. Second, it shows how administrative data can address some of the empirical challenges associated with the Labour Force Survey data.

The rest of the chapter is structured as follows. Section 1.2 describes the two datasets, their advantages and disadvantages and explains the procedure to build a quarterly panel from the administrative data; Section 1.3 explains the different expansions to the administrative data, checking the resulting unemployment rate against the Labour Force Survey official unemployment rate; Section 1.4 provides some further robustness checks; Section 1.5 shows how the expanded administrative data can help to evaluate and interpret two problems with Labour Force Survey data; Section 1.6 concludes.

## **1.2 Data**

This section explains the main characteristics of the two data sources I employ throughout this chapter: the Spanish Labour Force Survey (LFS thereafter), elaborated by the National Statistics Institute (INE in its Spanish acronym), and the Continuous Sample of Working Lives (MCVL thereafter), provided by the Spanish Social Security. It offers a comparison between the two and briefly explains how to structure the latter as a quarterly panel.

### **1.2.1 Description**

Official unemployment statistics come from the Spanish Labour Force Survey (LFS thereafter) which follows a representative sample of over 100,000 people for six consecutive quarters, with weights to reflect population characteristics. I will use these

weights when reporting stocks, as it corrects for the sampling errors - some groups are over-represented and some under-represented. The expand of the data I use is from 1987 to 2013. As other labour force surveys, it classifies workers by asking them to report their activities in the week of the interview - if they were employed or if they searched for a job, for example.<sup>1</sup> This allows the LFS to observe whether a worker is out of the labour force and why.<sup>2</sup> The LFS is thus structured as a panel.

However, many participants do not reply in all of the six quarters that they are interviewed, which leads to problems when calculating stocks and more so when building labour market transitions. These problems are partly corrected by adjusting the weights, which results in consistent stocks over time. Another complication arises from the changes the survey itself has undertaken over the years. In particular, two major changes in 2001 and 2005 affected how unemployment was recorded and produced breaks in some series.<sup>3</sup> These changes do not affect, by design, the stocks, but do alter labour market flows. In Section 1.5 I will explore these issues in more detail.

The Spanish Continuous Working Life Sample (MCVL thereafter) comprises the working histories of a 4% sample of the working population for the years 2004-2013. Similar datasets exist for Germany, Italy, Austria and the US, among other countries (see Tattara and Valentini (2010) for a table summary). The MCVL stands out as

---

<sup>1</sup>To classify labour market status of the population I use the variable *Type of Contract* for employees, *Current working situation* for the self-employed and the variable *AOI* for unemployed and out of the labour force individuals. This last variable encompasses the answer to other key variables (“Were you working this last week?”, “Are you looking for a job?”/“are you ready to work in the next 15 days?” and “What type of contract do you hold?”). This variable is the one used to build official unemployment rate series, which are reported in the EUROSTAT and OECD

<sup>2</sup>In particular, if the respondent is not employed nor looking for a job it asks her to declare the reason by choosing one of 9 possible answers. These include “studying”, “thinking they are not going to be able to find a job”, “caring for others” etc.

<sup>3</sup>The 2001 reform added the requirement for unemployed workers to be available for work in the next two weeks. This change caused a shift in the stock of unemployed in 2001. But the major change came in 2005, when the sample was altered to reflect the growing impact of migrant workers and an electronic way of carrying the survey from quarter was introduced.

being very accessible and big - there are more than 20 million observations in total as of 2013. It uniquely identifies workers and firms, allowing to observe job-to-job transitions and distinguish quits and lay-offs. It also keeps track of self-employment, something that is excluded in other datasets. The firm and worker identifiers allow to link the working histories panel to a yearly Income Tax complement, containing fiscal information about wages, other retributions in kind, unemployment or disability benefits, severance payments, and any flow of income between the firm and the worker (or the Social Security and the worker). For self-employed workers and firm CEOs, it contains declared profits, and although that information is highly susceptible to misreporting for tax avoidance purposes, it nevertheless provides an insight into self-employed earnings. These characteristics could in principle allow the dataset to be treated as a matched employer-employee data. However, as the unit of measurement is the worker and not the firm, it is very unlikely that we observe all of the workforce from the firms in the data.<sup>4</sup> Alternatively, it also contains a file that details the taxable income received by the worker in all of their previous spells. This is not the same as gross wages from the tax data, but as shown in Bonhomme and Hospido (2017) it can be adapted to study wage and earnings dynamics.

The raw MCVL data are not immediately ready to use as there are not designed for the purpose of research but for bookkeeping. There are some academic articles explaining how to clean and format the data (see Lapuerta (2010) or García Pérez (2008), for example). In particular García Pérez (2008) provides a comprehensive identification of the main problems when dealing with overlapping employment spells and censored unemployment spells. After implementing most of the cleaning and formatting procedures there is still the question of how to handle unemployed workers who are not registered within the Social Security. These periods appear as

---

<sup>4</sup>It could be argued that the sampling of firms is representative of the universe of firms in Spain, as the sample is representative of the worker side, it thus should be representative of firms too. As self employment is represented too, the coverage of small firms is good but few large firms are represented.

blanks, gaps between observed spells. This feature is common to other administrative datasets, but in Spain this issue is especially relevant because of the prevalence of very short and very long unemployment spells. These issues and how to deal with them are at the core of this chapter.

In principle it would be possible to use the retrospective information of the MCVL to build a panel earlier than 2004, as we have information on the complete working histories of workers, dating as far back as the 1960's. However, García Pérez (2008) warns against doing this kind of inference as the sample is representative of the year that it is collected from. That is, the 2005 file is representative of the working population of Spain in 2005. In the next years the sample adds new spells to adapt to demographic changes, but it does not “drop” any worker. The cases of workers dropping are either migration, transitioning out of the labour force or deceased.<sup>5</sup> Using the 2005 to do any inference on the labour market in 2000 would cause some relatively minor representativeness problems, as there was a substantial influx of migrants in the 2000-2005 period. But using the MCVL to look at the 1992 recessions would over-represent younger workers as the average age falls considerably. The sample size of the MCVL increased considerably in 2005 to account for better representation of different groups, so the MCVL of 2004 is not very well suited for study. For this reason, I follow García Pérez (2008) and only use the year 2005 onwards when building stocks, to make it comparable to the LFS. I will use the 2005 file to account for flows in 2004 as the sampling error of a year is not too significant. Finally, the potential accuracy gains that can derive from using only the final wave (2013) can outweigh the problems with representativeness for some applications. One of these exceptions is unemployment duration, as using the last year only provides with higher accuracy. Using instead all of the waves can result in more overlapping spells and discontinuities. Table 1.1 summarises the main characteristics of the two datasets.

---

<sup>5</sup>In this last case, the date is recorded in the MCVL.

Table 1.1: Data Comparison Table

	Labour Force Survey	Administrative data (MCVL)
Description	Rotating panel of quarterly interviews with a sample size of over 100,000. It is available since 1987.	A sample of about 1,000,000 job records of people with any sort of affiliation with the Social Security. It can constitute a panel since it follows most of the same people over the period 2004-2012.
Advantages	<ul style="list-style-type: none"> <li>* Detailed and accurate information for personal characteristics (such as education).</li> <li>* It has the potential to track labour status changes that are made out of the scope of administrative records (first job seekers, informal market jobs, inactive workers)</li> </ul>	<ul style="list-style-type: none"> <li>* Firm and worker identifiers allow for the study of job-to-job transitions (rarely available in the LFS).</li> <li>* Very accurate information on employment spells, with precise dates of entry and exit into and out of jobs/unemployment</li> <li>* Can be matched to a fiscal dataset for wage/benefit information.</li> <li>* Consistent through time.</li> </ul>
Disadvantages	<ul style="list-style-type: none"> <li>* Fails to capture short term jobs and some very short unemployment spells due to it being a <i>quarterly</i> dataset.</li> <li>* There are important series breaks in 1992, 1999, 2001 and 2005.</li> </ul>	<ul style="list-style-type: none"> <li>* It can't track anyone who has no formal relationship with the Social Security. As such, it serves poorly for tracking people out of the labour force.</li> <li>* For the same reason, it is also unable to track informal market activities.</li> </ul>

## 1.2.2 Prepare the MCVL for use

In order to work with the MCVL to analyse job market variables it is necessary to at least establish a reference variable for labour market status and treat some simultaneous spells. If we also want to make meaningful comparisons with the LFS, we need to build a panel with one observation per quarter. Laborda (2013) for example uses this approach. Appendix A provides a detailed guide of the formatting process. Here I explain the main steps of the procedure to make the MCVL ready for use in research:

1. Classify the labour status of the worker in each spell
2. Modify overlapping spells and extend censored spells
3. Build a panel by selecting one spells per period of time

### Labour Status of the worker

The aim of the first step is to create a variable that classifies workers in four categories: *self-employment*, *working with a permanent (open-ended) contract*, *working with a temporary contract* and *unemployed*. It is important to separate both kinds of contract because their dynamics are very different, with temporary contracts accounting for 90% of all job creation<sup>6</sup> and most of the flows in and out of unemployment.

The only category missing is *out of the labour force*. The Social Security does not provide sufficient information to judge whether someone is participating or not in the labour market. In order to keep their benefits, unemployed workers are formally required to: prove they are actively searching for a job, attend job interviews and not reject job offers. The Employment Centre monitors workers at least each month

---

<sup>6</sup>Temporary contracts who expire and are not renewed do not incur into dismissal transfers, which is the case of termination of open-ended contracts. These contracts always end in dismissal, quit or retirement.



upon receiving the payments. A priori, I treat this as sufficient proof of unemployment.<sup>7</sup> Retired workers are not part of the main sample, and their information is in another linkable dataset. They constitute, according to García-Pérez, 27% of the total number of observations in 2007, which is far from the 40% of inactive workers that the LFS reports for that year. The remaining 13% should be individuals that are too young or have never participated in the labour force. As these groups are excluded from the sample, the remaining option is to count periods in which the Social Security has no information on the worker as out of the labour force. This is the initial treatment I give them, as in previous studies. Later on I relax this assumption.

Three variables contain all the information needed to classify workers in the same working status as in the LFS:

1. *Type of Labour Relationship* (TRL) codes the different links each worker has with the Social Security – working, receiving unemployment benefits. This way I separate unemployed workers.
2. *Contract Type* contains the code for each type labour contract.<sup>8</sup> There are 557 kinds of different contracts in the registry, but most of them are “legacy contracts” that do not exist in the present. Most temporary contracts are grouped under the 400s numerical codes while regular contracts are coded in the 100s. This way I distinguish between temporary (which have a specific termination date) or permanent (open-ended) contracts.<sup>9</sup>

---

<sup>7</sup>This may not be the case if monitoring fails, for example if an unemployed worker that lies about searching for a job. This is acknowledged by most authors (see for example García Pérez (2008), Lapuerta (2010) and Arranz et al. (2011)). García-Pérez points out that most of these dropouts from the sample correspond to women and young people, who are less attached to the labour market.

<sup>8</sup>There are some kinds of contract that don't exist any more - usually contracts with some kind of temporary subsidy created in the 1990's. These are not relevant for the present study as I focus in the 2005-2013 period.

<sup>9</sup>In the particular case of discontinuous workers (those who work only on specific periods of time every year) I treat them as permanent, as they are subject to firing costs and have no pre established termination date.

3. *Contribution Class* allows for the identification of self-employed workers, as they have a different arrangement with the Social Security. These correspond to variable values 500-600.<sup>10</sup>

Using those three variables suffices to classify most observations, but there are special cases: the unemployed close to retirement that choose to pay their contributions to the social security as if they were employed to boost their pension, discontinuous and seasonal workers who get a compensation in between working seasons or students that receive benefits under apprenticeship contracts. The treatment for these specific cases is left for appendix A.

### **Clean overlaps and extend unemployment spells**

After classifying the state of the worker, it can still be the case that a worker is classified in different status at the same time. García Pérez (2008) recommends to drop overlapping spells where the worker is simultaneously employed in more than one firm<sup>11</sup> or when the beginning and the end of the spell overlap for a few days. It can be the case where employed or unemployed individuals are receiving some form of complementary benefit during their current spell. For example, because of an incapacitating illness or a widowhood pension. In these cases the best approach is to keep the employment spell only, and to merge some of these overlapping spells in unemployment.<sup>12</sup>

The second step can be omitted if the researcher wants to use the MCVL definition of unemployment. García Pérez (2008) recommends to add to unemployment between two employment spells the missing days before and after the unemployment

---

<sup>10</sup>There are some special categories for domestic workers, agriculture workers, farmers and sailors. I select those who are self-employed in these special regimes.

<sup>11</sup>These cases are rare in Spain and mostly refer to part-time jobs. However some employees of the Catholic Church are recorded to have two full-time jobs: one for their religious duties and one for their other dealings - teaching for example.

<sup>12</sup>The Social Security records pensions separately, but the pension file can be easily merged with the main working records file using the individual identifier.

spell if these are not recorded, especially in the case of days between the end of the job and the beginning of the unemployment spell. These are likely due to administrative delays and should be added. In Section 1.3 I analyse the different approaches to unemployment spells in more detail.

## **Building the panel**

Once each worker has a unique observation for each quarter, we can proceed to build a panel by only keeping the observations that relate to each quarter of interest. The easiest approach would be to take the 2013 file and use its retrospective information. However, as each year file is representative for the population in that particular year, using the 2013 retrospectively will lead to a bias in favour of younger workers in previous years, as discussed before. The second main reason is that is more practical for handling the data: to mimic the LFS we would need to sample a spell that lasts beyond the year as separate, different spells. That is, year frequencies allow for the building of an observation per year per worker, with details of the states in the different quarters. In this way, each worker ends with an entry for each year she is in the sample. Using a unique file would require the use of many more auxiliary variables.

To build the panel I first establish a point in time at which I will evaluate people's working states: the two weeks starting the 1st of January, 1st of April, 1st of July and 1st of October, which coincide with the start of the year's quarters. Because some jobs may start after that date, I also consider all the spells in the following two weeks, until the 15th of each month. The labour status variable will reflect the relationship each individual have on those weeks: self-employed, open-ended employee, temporary employee or unemployed.<sup>13</sup> Laborda (2013) uses the fifth week or

---

<sup>13</sup>This corrects for sort periods in which the worker may not be in either state. For example, it is likely that many jobs will start on the 7th of January instead of the 1st, due to Christmas holidays in Spain ending on the 6th.

each quarter, as the LFS takes place around that date. Most job contracts tend to start on the first day of the month, so the first weeks of the quarter seem a natural choice. For workers that have more than one spell in the same two-week period, I give priority to the longest spell. That is, if a worker starts the two-week observation window unemployed but end with a job that lasts for one more quarter, I count her as employed on that quarter. If there is a tie (0.03% of the total number of observations) I chose employment over unemployment, and self-employment over employment. This is because some jobs (especially part-time) can be complemented with unemployment benefits, but that doesn't mean the worker is unemployed.

## 1.3 Methodological Check

The challenge is thus how to treat unemployment spells in the MCVL. In the second step of the procedure outlined in Section 1.2.2, the researcher needs to take a stance on what to consider as unemployment, or alternatively as in Alvarez et al. (2015) consider all the gaps between employment spells as non-employment. In this section I argue that some extensions that allow us to consider most of these gaps as unemployment spells, providing a series of methodological checks against the LFS unemployment series. The motivation of these expansions is to match the unemployment rates from both datasets, which differ considerably after 2008.

### 1.3.1 The unemployment gap

The LFS and the MCVL have a different number of observations (an average of 108,136 in the LFS<sup>14</sup> and 678,183 in the MCVL) so in order compare the stocks I express them as shares of the labour force thereafter.

---

<sup>14</sup>The weighted labour force survey has a mean of 31,360,266 observations per period – which amounts to the total population of Spain.

Figure 1.1 shows the main discrepancy between the MCVL and the LFS: the unemployment rate. This disparity reaches 10 percentage points by the second quarter of 2013. The differences persist when unemployment rates are compared by gender and sex in figures 1.2 and 1.3. The two unemployment rates are closer for men than for women. For young workers between 20 and 30 years of age, the MCVL unemployment is half of the LFS. For prime aged workers (30-50) the differences are more modest. The differences are starker for older workers, whose trend and level differs between the MCVL and the LFS.

The main source of the differences comes from the different definitions of unemployment they use:

- The LFS considers a person unemployed if: (1) they are not currently employed (2) are actively looking for a job and (3) they are ready to start working within the next 15 days.
- The MCVL considers a person unemployed if: (1) she has been in the social security system before (had a previous job) and (2) is receiving unemployment benefits.

This means the MCVL excludes all unemployed whose benefits have expired. The Spanish Social Security does not provide any other benefit for those who exhaust their unemployment compensation, so all unemployed beyond 2 years<sup>15</sup> *disappear* from the registry. As long term unemployment reached 50% of total unemployment by the end of 2013, this is the principal potential source of disagreement. The first expansion deals with this issue by extending observed spells until the start of the next job or the end of the sample.

---

<sup>15</sup>this limit is expanded to an absolute maximum of 4 years for workers with family obligations or other special circumstances

Figure 1.1: Unemployment rate in Spain

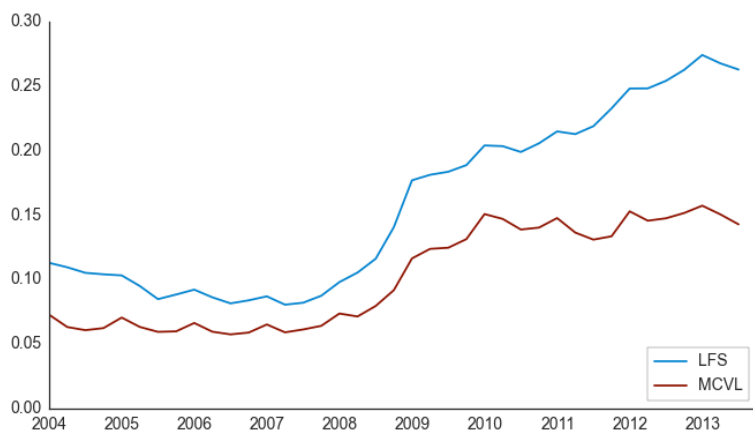


Figure 1.2: Unemployment rate by sex

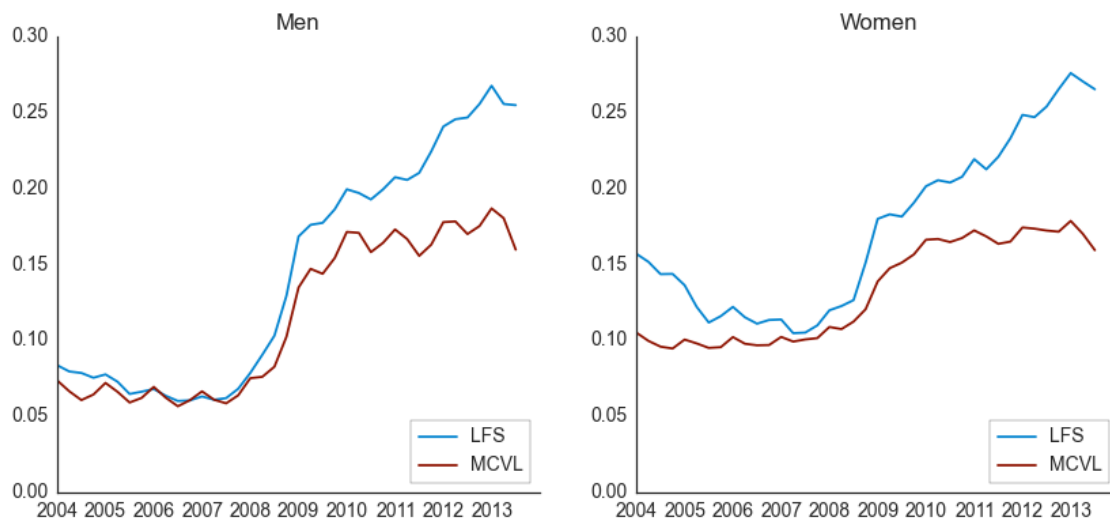
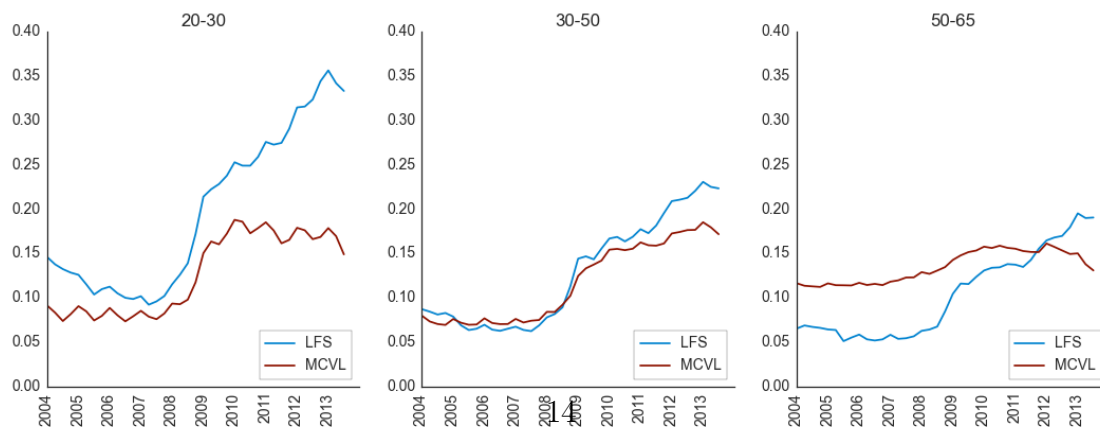


Figure 1.3: Unemployment rate by age



The Social Security also excludes all unemployed without the right to claim unemployment compensation (who have been employed for less than a year in the last 4 years) and unemployed that have not held a job yet. The second expansion aims to recover these spells (usually related to young workers in short lived temporary contracts) by adding gaps between employment spells of those without the right to claim.

Finally the Social Security may be counting as unemployed some inactive workers by the definition of the LFS (not actively searching for a job and/or not ready to work in the next 15 days). This would imply the MCVL has a bias upwards with respect to the LFS. But this is not what we observe in the data, except for older workers.<sup>16</sup>

### 1.3.2 Closing the unemployment gap: LTU expansion

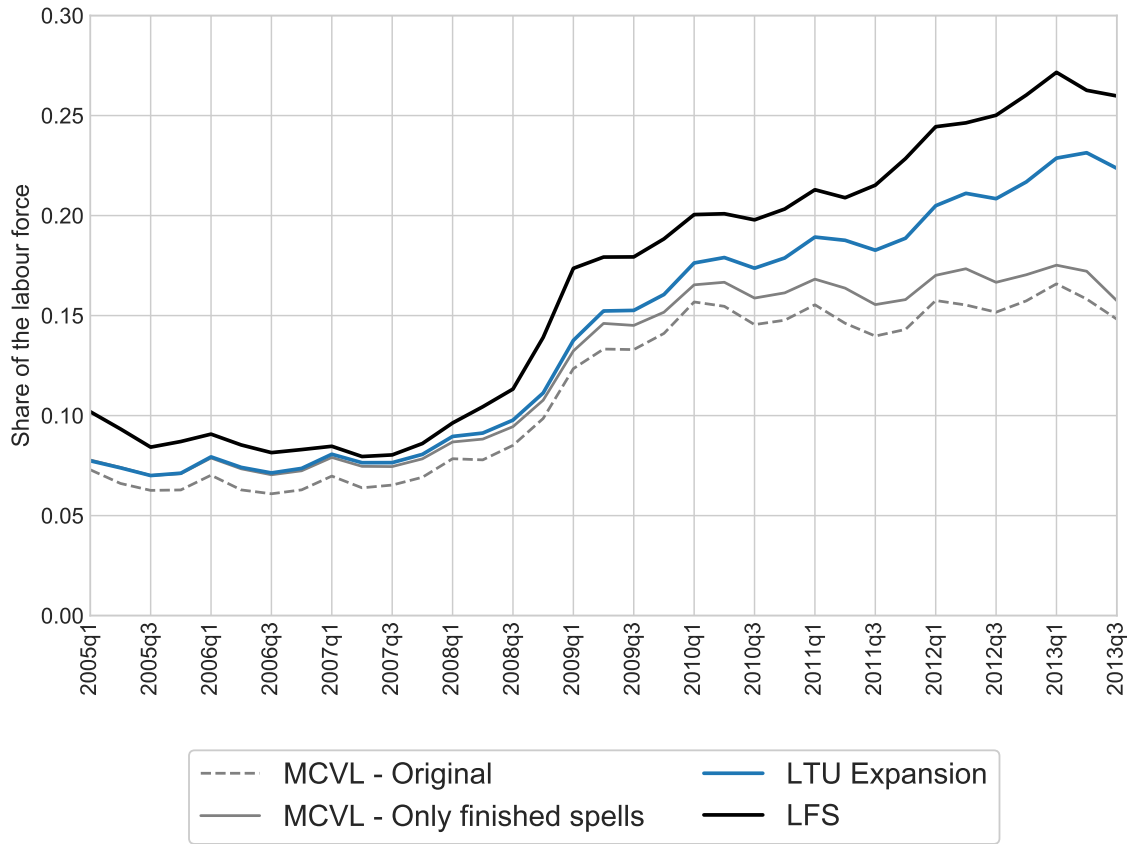
Given the importance of long term unemployment in the last years in Spain, it is natural to start by adding the days elapsed between the recorded end of the unemployment spell and the start of the next job. This is already suggested by García Pérez (2008) as a necessary treatment to work with the MCVL. This extension is easy to implement but has a shortcoming: many of the long term unemployed have not found a job by the end of the sample. This is reflected on the small difference between *Original* and *Only finished spells* unemployment rates in figure 1.4: the difference in trend and level of unemployment widens from 2009 onwards. In fact, comparing it with the original MCVL series, it barely makes a difference.

The *LTU Expansion* adds all the unfinished unemployment spells by the end of 2013, as well as extending the duration of registered unemployment spells between jobs as before. After this expansion both trend and level are very close to the LFS, as

---

<sup>16</sup>We can observe when an individual is receiving unemployment subsidies immediately before retirement in the MCVL, so they are excluded from the plot.

Figure 1.4: Unemployment rates - LFS and LTU expanded Social Security



shown in figure 1.4. The expanded series is still below the LFS by 3.7-2.5 percentage points by 2013. This shows us that if we want to match the trend of the LFS unemployment rate, we need to add unfinished spells. How many depends on the sample size. 2013 was the worst year of the recession in Spain, so many unfinished spells are to be expected. But if a researcher has access to further years, most of the difference may be captured by adding the finished unemployment spells only. I also condition the extension of unfinished unemployment spells on being at most of 2 years.<sup>17</sup>

<sup>17</sup>Robustness checks can be used to asses if the 2 year maximum duration is a good threshold.



### 1.3.3 Closing the unemployment gap: STU expansion

Other kind of unemployment the MCVL is missing is people without the right to claim. This will be the case of:

- Quits to unemployment. Quits do not have the right to unemployment compensation.
- New entrants to the labour market (with no previous employment record)
- Temporary workers with employment spells below the minimum requirement - less than a year of tenure in the last 4 years.
- Self-employed workers who have no right to unemployment compensation.

Notice that all of these cases are not the main source of discrepancy between the LFS and the Social Security, as the the first expansion is already very close to the LFS. Underestimating long term unemployment a bigger issue, but capturing short term spells is important for matching youth unemployment rates. Therefore I refer to the resulting expanded unemployment series as the *Short Term Unemployment* (STU) expansion.

To identify these spells, I chose to include all gaps between employment spells that last at least 15 days and at least one of the following conditions:

- The worker quit her last job.
- The worker was self-employed in her last spell.
- The worker does not have enough contribution periods to be eligible.<sup>18 19)</sup>

---

<sup>18</sup>The threshold is less than 360 days of employment, according to Spanish legislation.

<sup>19</sup>It is worth noting that the law in Spain does not allow to claim benefits that were not consume in the last unemployment spell if the worker wants the past employment spell to count for future benefits. For example, say a worker has 3 months left of UB, and finds a 6 month job; after the job ends, she can chose to claim the 3 missing months from before or the 2 months she has accumulated with the last job. She can't have both.

Table 1.1: STU Expansion spells, by type

	Quit	Self-employment	No right to UI
Total	151,461	91,972	551,272
Percentage	19.06	11.57	69.37

Notes: Elaborated with data from the MCVL, waves 2005-2013

In addition to these restrictions I add the requirement that the worker is not to be recalled to work on the same firm. The reason for this is that there is a good chance that the worker knew that she was going to be called back and thus had no incentives to search. This is particularly important because employers could use these tactic to extend the maximum duration of temporary contracts. That is, instead of renewing the worker beyond the two month period, the firm asks the worker to take a period of leave and then return.

The conditions imposed make it very unlikely that a person detached from the labour market would qualified for an unemployment extension. Table 1.1 shows the different spells that are added in the STU expansion by each case. Most of them do not have the right to claim unemployment benefits, but a non-trivial amount also come from self-employment and quits. Figure 1.5 shows the histogram and empirical CDF of the ages of the unemployed at the time of the start of their unemployment spell, broken down by expansion. Unemployed individuals from the STU expansion are overwhelmingly younger, with 80% of them under 40. In terms of what was their previous spell, figure 1.6 shows that they are more likely to come from temporary contracts than in the LTU expansion, and their next spell is also more likely to be another temporary job. This gap increases if we consider that virtually all unemployed workers coming from self-employment are in the STU expansion. If we exclude self-employment, 86% of all previous spells in the STU addition are temporary jobs, while the LTU and the original only have 70% and 74% respectively.

Figure 1.5: Age Distribution by Expansion

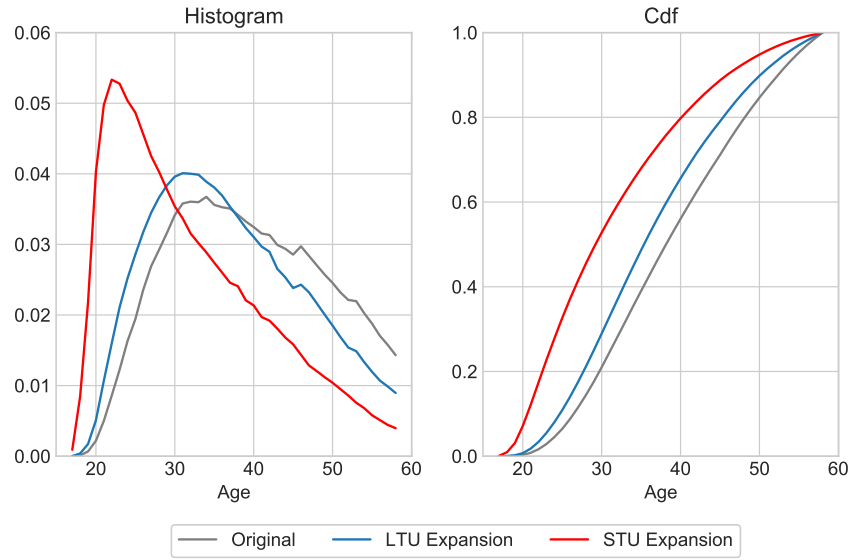
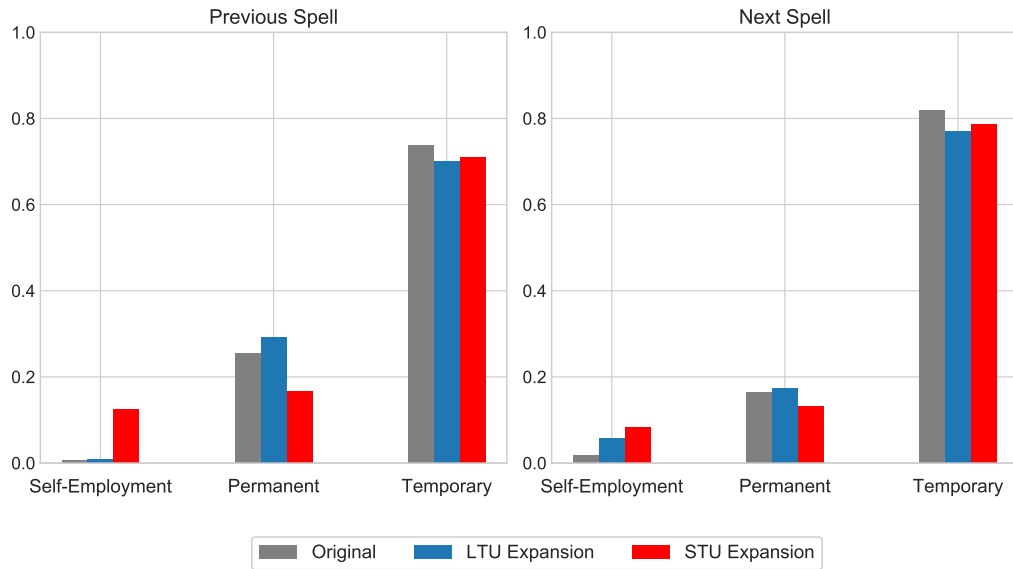


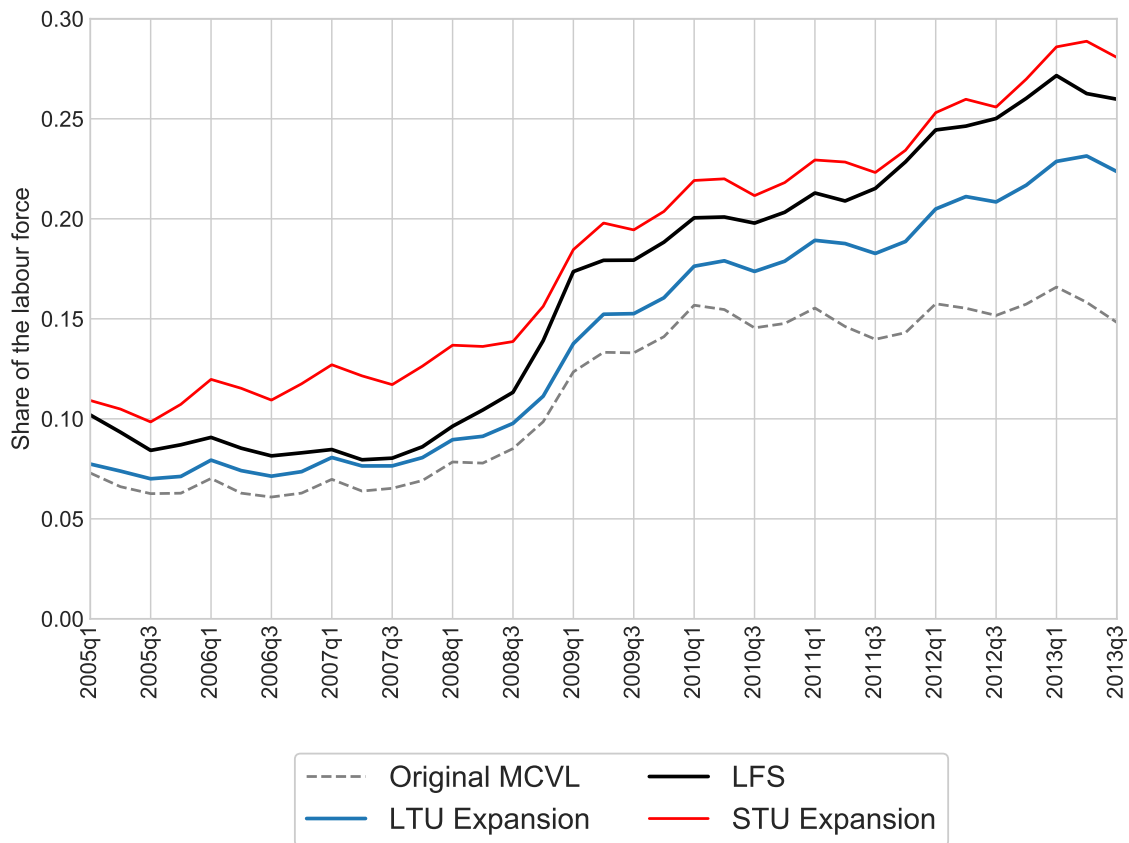
Figure 1.6: Spells before and After Unemployment



Appendix A.2 analyses these spells in more detail.

After adding these spells, the MCVL unemployment rate gets closer to the LFS after 2009, as figure 1.7 shows. There is still an overestimation of unemployment before 2009, but as for the end of 2013 the differences are small. It is not surprising that the STU expansion adds more unemployment relative to the Long-Term expansion in the 2005-2008 period. These years coincide with the construction boom and the highest rates of temporary contracts over total employment. In the following years the gap is reduced as long term unemployment increases its incidence. The trend is similar to the Long-Term expansion and the LFS.

Figure 1.7: Unemployment rates - STU Expansion



We can gain some insights into the difference between the two expansions by looking at the unemployment rates broken down by gender (figure 1.8) and age (figure 1.9). By gender, the Short-Gaps expansion brings the MCVL closer to the LFS. This is expected as women are more commonly employed in the services sector, where temporary contracts are very common. It is less successful for men, in particular before 2008. This again can relate to the use of temporary contracts among construction workers, but notice as well that even the original MCVL overestimates unemployment in this period. This suggests that men are less likely to report to be searching for a job when claiming benefits. The overestimation of unemployment persists even after onset of the recession.

By age things are also clear: the STU expansion helps reconcile the unemployment rates of younger workers, in a way the Long-Term expansion is not able to match. There is a small positive gap in the 2006-2008 years, again likely driven by males on temporary contracts. For middle-aged workers the differences mirror those in figure 1.7, as it is higher than the LFS (in particular for the 2005-2008 period). Here the Long Term expansion performs arguably better. For older workers the STU expansion barely makes a difference over the Long-Term one. Still both offer a more coherent picture than the original MCVL series. For this age group, notice the overestimation of unemployment even for the original MCVL in the 2005-2008 period. Here it becomes clear that the definition of unemployment matters: most of those who report not to be looking for a job “because they think they are not going to find one” are in this age group. Although they are receiving unemployment benefits and would like to work, they report a low willingness to search, which makes them inactive in the LFS. The big disparity in their unemployment rates does not affect however the overall unemployment rate, as the share of older workers in total unemployment is small. Both expansions improve on the trend of the original MCVL, especially for middle-age workers and the LTU expansion.

Figure 1.8: Unemployment rates by gender

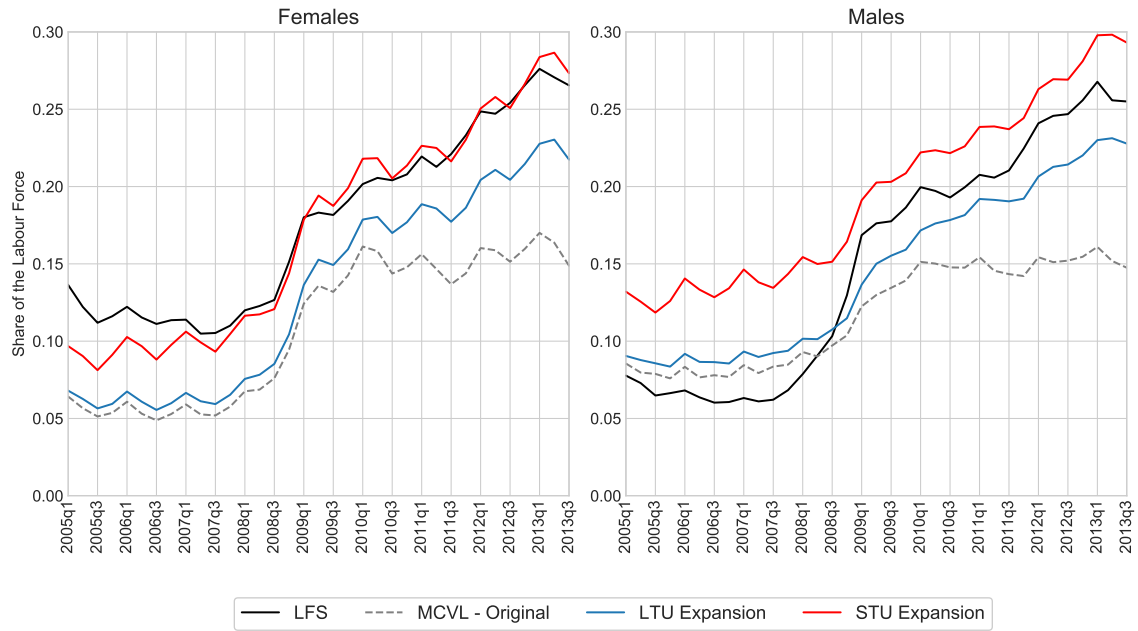
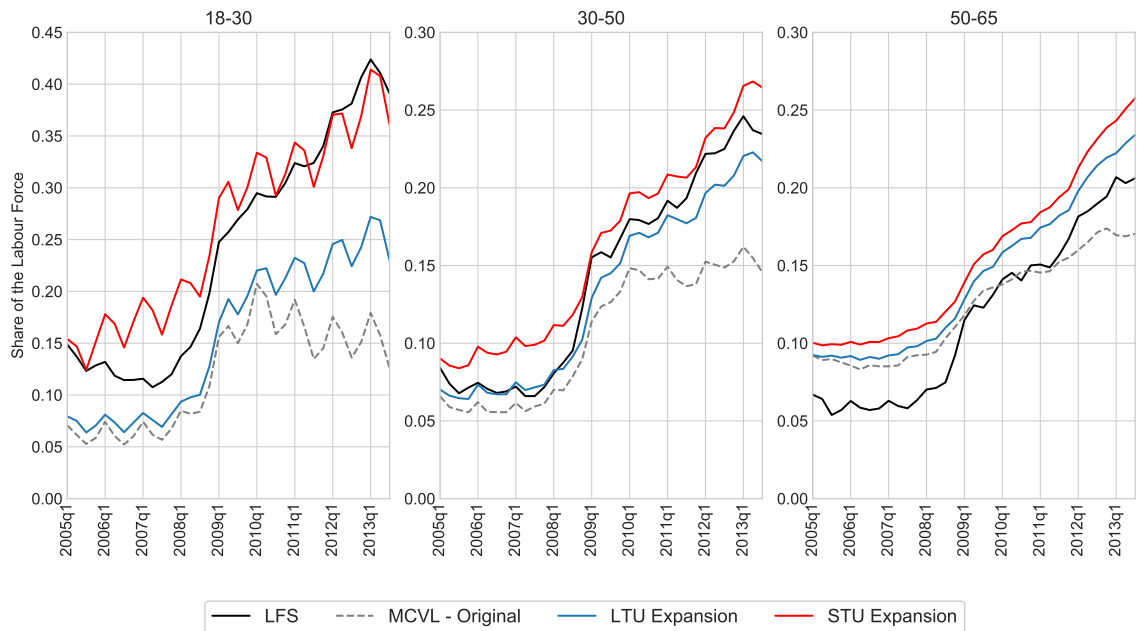


Figure 1.9: Unemployment rates by age group



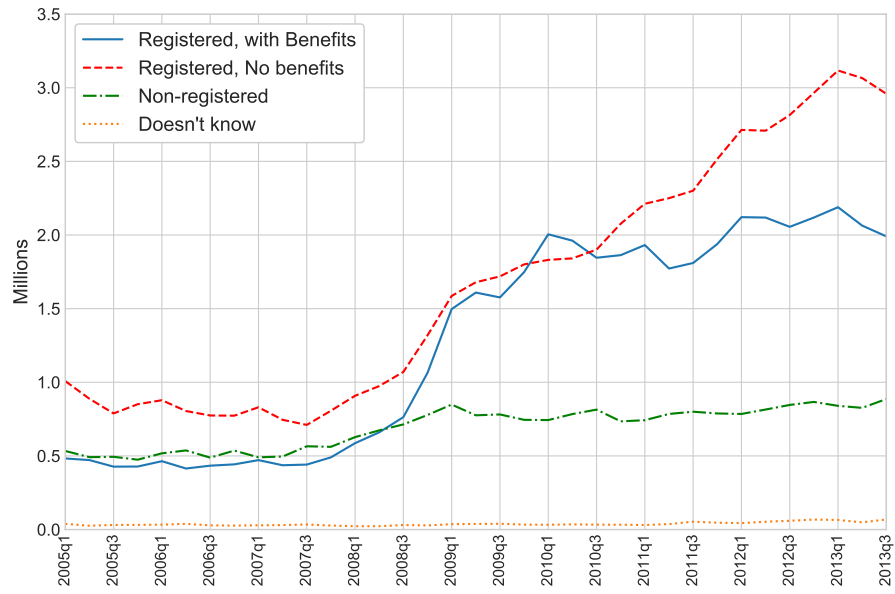
## 1.4 Further Robustness checks

In the previous section I created different unemployment series using the information from the MCVL and institutional setting. I then checked the resulting stocks against the LFS. This section provides a further check comparing the extent of self-reported unemployment without benefits in the LFS to the expansions of the MCVL. This check is interesting because it shows that as it was possible to go from the MCVL to LFS definition of unemployment, the reverse path can also be traced up to a certain extent. The methodological differences between the two datasets mean that the match is not perfect, but nevertheless very close.

The fact that the LTU expansion managed to bring the MCVL closer to the LFS was derived directly from the definition of unemployment in the MCVL. But we can use the information from the LFS to check if there is any clear trend in benefit expiration that would explain why the LTU expansion gets closer to the data. This check can be done by looking at a variable in the LFS called “Relationship with the Employment Office”, that asks individuals if they are registered as unemployed and if they are receiving benefits. This variable thus classifies individuals in four categories, ‘Registered, with Benefits’, ‘Registered, No Benefits’, ‘Non-registered’ and ‘Doesn’t know’. The first case will be recorded in the MCVL, while the second should not be recorded. These unemployed without benefits are the ones that the LTU correction is targeting. The third case corresponds to job seekers that are not registered with the public Employment Office. These cases will not be recorded by the Social Security, and are the ones the STU expansion seeks to recover.

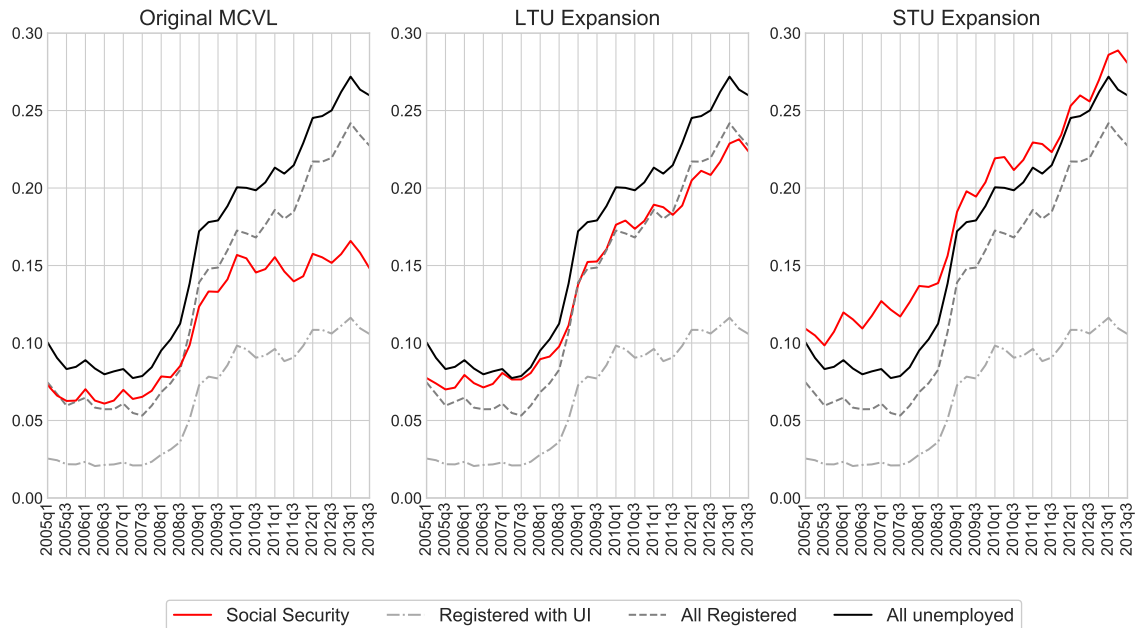
Figure 1.10 shows the evolution of these stocks (in millions) in the 2005-2013 period. The *‘Registered, with Benefits’* line looks very similar to the original MCVL series, which highlights again that these are the only unemployed captured in it. The stock of *‘Registered, No Benefits’* on the other hand looks more similar to the

Figure 1.10: Relationship with the Employment Office



Source: LFS

Figure 1.11: Alternative unemployment rate series





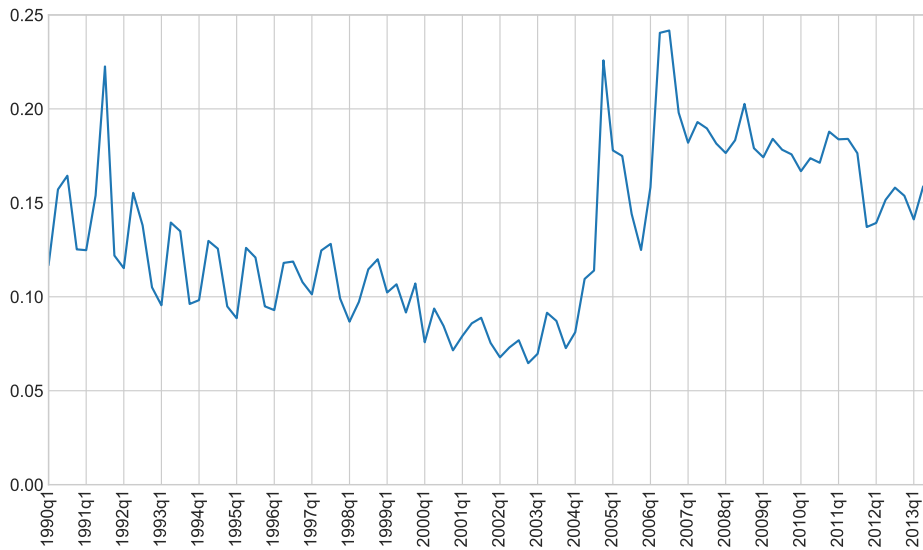
expanded series, with a big increase after 2008. The *'Non-registered, No Benefits'* stocks do not change substantially in the period, increasing slightly after 2008. These trends are reassuring, as they correspond to the different expansions.

In figure 1.11 I use this variable and the labour stocks from the LFS to build alternative measurements of unemployment. The first panel shows the the original MCVL unemployment series (in red) has very similar trend to the LFS series were only the unemployed who are receiving benefits are considered. However, the MCVL is higher and corresponds to only the registered unemployed before 2005. The second panel shows that the LTU expansion matches closely with the case where only those registered in the Social Security (with or without benefits) are considered, in particular after 2008. This provides additional evidence that considering only those who are registered in the employment office is not enough to account for unemployment after 2008. The last panel shows how the STU expansion achieves a closer unemployment rate to the LFS when all of the unemployed are considered after the recession, but overestimates unemployment before. As discussed in Section 1.3.3, this likely reflects that the MCVL captures more frictional unemployment than the LFS, which was more relevant before 2005 than after.

## 1.5 Using the MCVL to enrich the LFS

As the unemployment rate from the expanded MCVL is similar to the LFS, we can use this to compare other labour market magnitudes. In this section I will be focusing on labour market flows, which are problematic in the LFS for two reasons: (1) respondents not replying on consecutive interviews, which affects the flows and (2) changes in the design of the survey which changes workers labour market status, in particular in 2005. The aim of this section is to show how the MCVL can clarify these issues.

Figure 1.12: Attrition of Unemployment Stocks in the LFS



### 1.5.1 Attrition and Labour Market Flows

The LFS is a rotating, panel, such that each household is interviewed in 6 consecutive quarters. I define *attrition* as a respondent who is not in her last interview fails to report on the subsequent interview. The size of the attrition bias has not been constant over time nor does it affect all individuals in the same way. Figure 1.12 shows the share of respondents who are not in their last interview and report being unemployed any given quarter, but do not respond to the survey in the next quarter. For example, about 8% of all individuals reporting being unemployed in the 2000-2005 period do not respond in the next quarter. After 2005, that number shoots up to over 15%, reaching 20% in some quarters.

The LFS corrects for this problem by changing the weights of the observations each quarter and introducing more people in the sample. This makes stocks consistent over time. However if we want to calculate labour market transitions the weights do not solve the problem. This problem is common among labour force surveys, and

usually resolved as in Silva and Vázquez-Grenno (2013) and Elsby et al. (2015): Taking the stocks as given, but treating the transition rates as biased because of attrition. Then we can calculate the transition rates that are consistent with the evolution of the stocks.

In all of these cases the stocks are given by the sum of flows in each quarter. For example, consider the transition from state  $X$  to  $Y$  as the number of observed individual transitions between  $X$  and  $Y$ , divided by the sum of all individual transitions starting from  $X$ , as equation 1.1 shows:<sup>20</sup>

$$\lambda_{t,flows}^{XY} = \frac{Y_{t+1}|X_t}{\sum_Y Y_{t+1}|X_t} \quad (1.1)$$

Assume that there is attrition in this data, but that it does not affect the transitions from  $X$  to  $Y$ , but a number of remainders ( $X_{t+1}|X_t$ ) are missing. Then the denominator would be *lower* than what it should be, as the non-respondents are taken out of the sample. Consider instead the the transition rate defined as in equation 1.2 below: number of observed individual transitions between  $X$  and  $Y$ , divided by the number of observed individuals in state  $X$ .

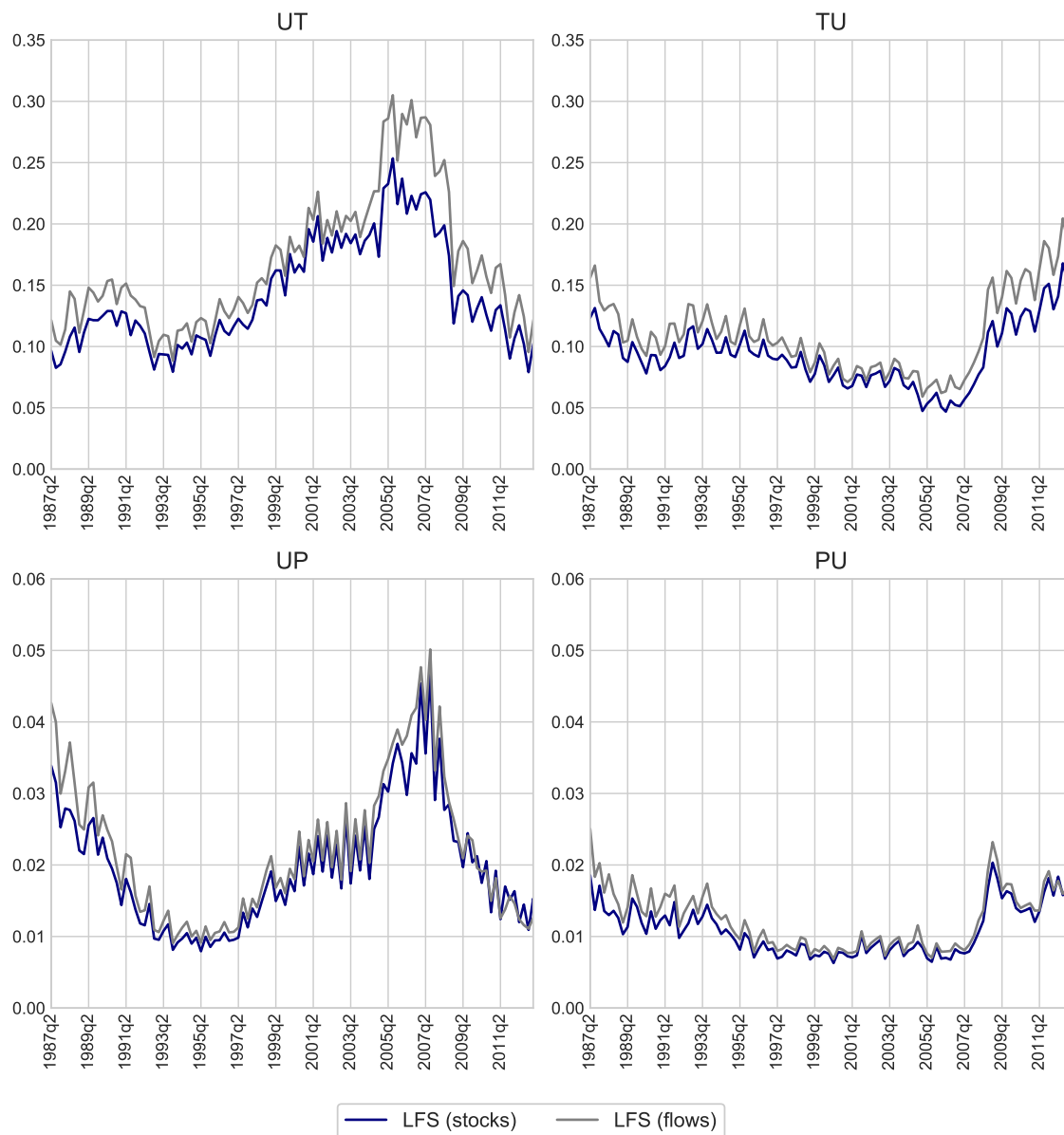
$$\lambda_{t,stocks}^{XY} = \frac{Y_{t+1}|X_t}{X_t} \quad (1.2)$$

In this way the transition rate would be consistent with the data. In practice, attrition can affect all of the rates out of state  $X$ , so the resulting bias of  $\lambda_{t,flows}^{XY}$  is ambiguous. We can consider the case of  $\lambda_{t,stocks}^{XY}$  as the extreme case when all of the attrition comes from stayers. That is, the non-respondents are not transiting to any other state in the next quarter. Figure 1.13 shows the evolution of  $\lambda_{t,flows}^{XY}$  and  $\lambda_{t,stocks}^{XY}$  from 1987 to 2013. There is not much difference between the two except in the flows between unemployment and temporary contracts. Here the gap is very no-

---

<sup>20</sup>I define transitions rates *forward* - from one quarter to the next. The literature tends to use the *backwards* approach - transitions from the previous quarter to the present. This distinction does not matter for results.

Figure 1.13: Labour Market Flows in LFS



ticeable in the 2005-2008 period, which coincides with the attrition ‘jump’ in figure 1.12. The gap is also noticeable for the temporary to unemployment (TU) rate after 2008. The MCVL does not suffer from this bias, as we can observe the changes in labour status of workers with more precision – up to daily frequency. The definitions of unemployment are different, as discussed, but given that the expansions get them closer we can compare the resulting transition rates to the LFS. As the MCVL does not suffer from attrition issues, comparing the LFS and the MCVL can give us some insight into the source of the discrepancies in the LFS flows due to attrition.

Figures 1.14 - 1.15 compare the flows resulting for the LFS to the MCVL. *LFS (flows)* shows the transition rates from the LFS calculated as in equation 1.1 (the denominator being the sum of transitions) while *LFS (stocks)* shows it as in equation 1.2 (the denominator being the stock).<sup>21</sup> The blue lines correspond to the LTU expansion of the MCVL and the red line to the LTU expansion. Given the increase in attrition of unemployed workers in 2005, I have taken back the MCVL to 2003 to have a larger window for comparison.<sup>22</sup> In general, the level and trend of the flows is close between the two datasets. The flows from the MCVL have higher seasonal variation because it captures short employment spells that the LFS can’t. Conversely, the LFS series are smoother because of its quarterly frequency. Notice that the flows version of the LFS is always higher than the stocks version, which would be consistent with the non-respondents being unemployed the next quarter as well.

The left panel of figure 1.14 shows that there are important differences between the stocks and flows version of the LFS in the 2004-2008 period. This could be indicative that the *UT* flow is not affected by attrition, but the denominator is. In

---

<sup>21</sup>When calculating the stock, I naturally exclude those who are in their last interview, as they would not reply in the next quarter because they are out of the sample.

<sup>22</sup>The observations from before 2005 are taken from the 2005 file, so there might be some representatively issues. I’m assuming these are not substantial for 2 years earlier in the sample.

Figure 1.14: Flows out of unemployment

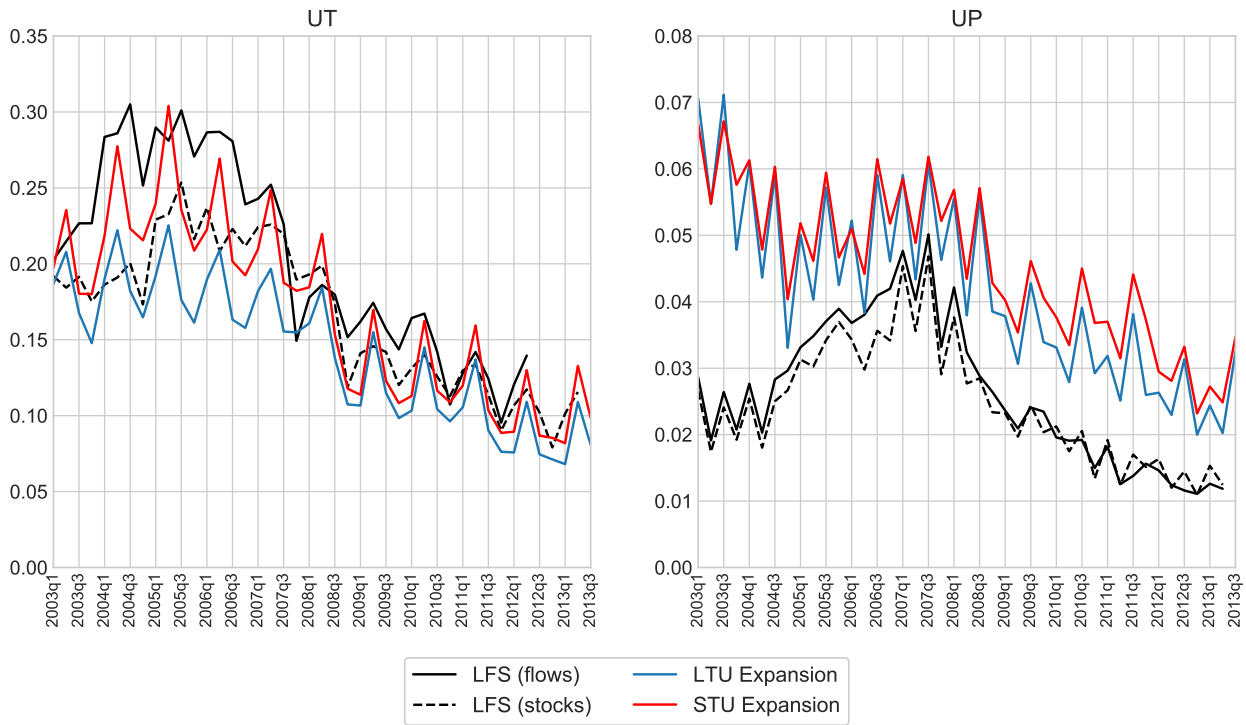
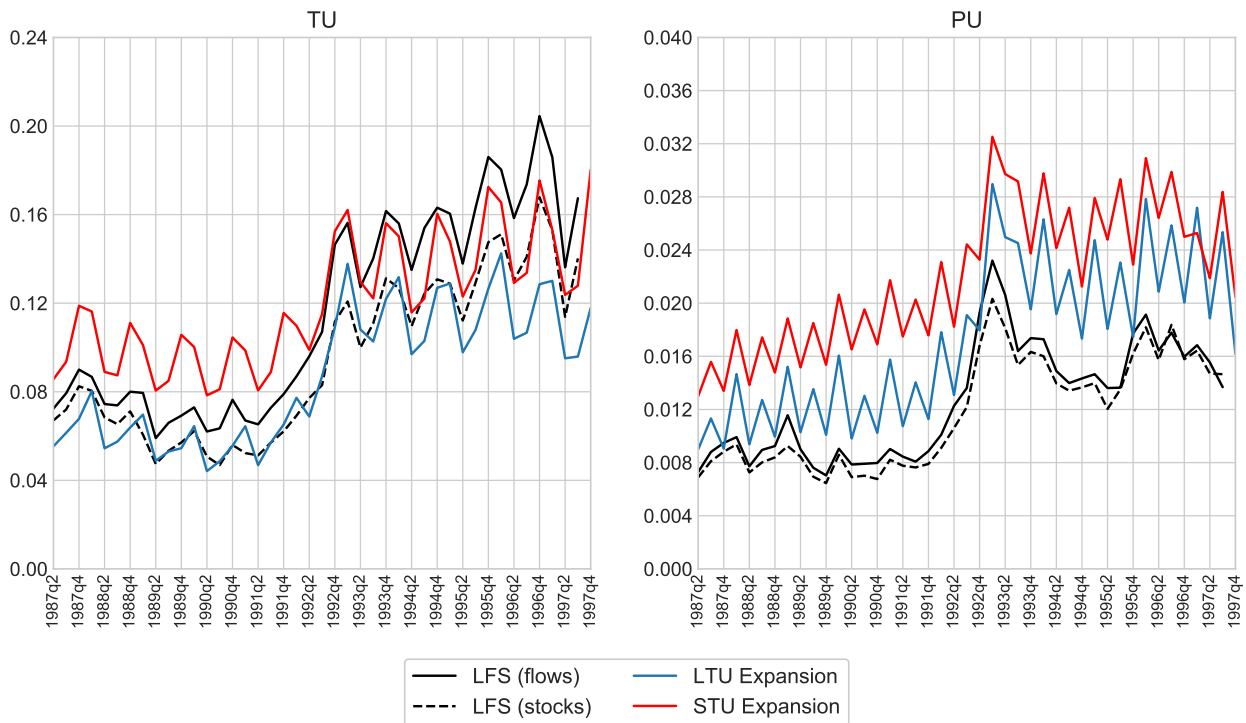


Figure 1.15: Flows into unemployment



this way, the non-respondents are not likely going to come from workers that take in temporary jobs. This discrepancy coincides with a higher discrepancy between the LTU and STU expansions of the MCVL. The differences between the expansions are clear: the STU expansion is capturing more movements into and out of temporary contracts, corresponding to more frictional unemployment. After 2008 however all series converge. This suggests that the attrition bias is partly driven by the unemployed without right to claim benefits flowing in and out of temporary contracts, which are not adequately captured in the LFS due to its frequency.

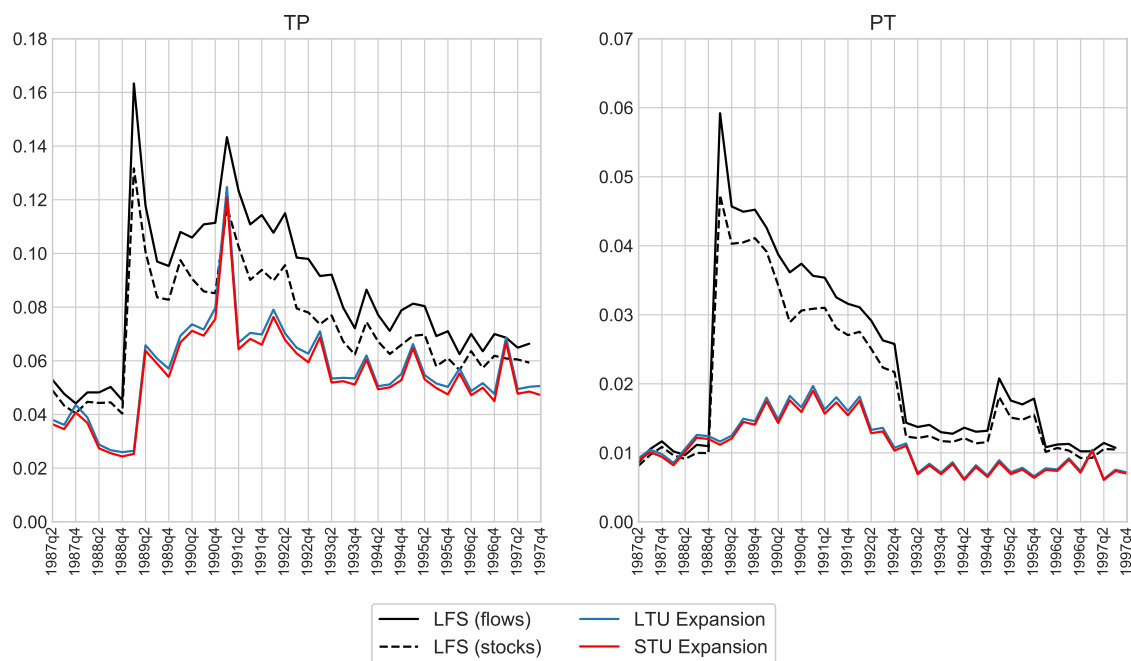
The right panel of figure 1.14 shows that the unemployment to permanent flows are higher in the MCVL, but the differences are small - notice that the scale is only from 0 to 8%. The divergence of the MCVL before 2004 can be partly explained by contract modifications adjustment not being marked before 2005 and it is a reminder that the data can't be take retrospectively without major problems. As for figure 1.15, the same conclusions carry over: the STU expansion is adding some short spells from temporary contracts that otherwise would be taken as job-to-job transition. The LFS does not capture well these quick changes and has a tendency to smooth them out, so both LFS series are below the STU expansion. This changes after the recession, as the volume of turnovers increases considerably. The LTU expansion matches quite closely the LFS (stocks) series.

In sum, the expanded MCVL is very close to the LFS, and helps explain that the attrition bias of the LFS is in part related to short-term unemployment spells coming from temporary contracts. Further research would have to confirm this, but the comparison with the MCVL points in this direction.

## 1.5.2 Changes in Survey Design

Attrition is not the only challenge when computing flows with the LFS: changes in the structure of the interview have also cause several discontinuities. These breaks are not present in stocks, because the National Institute of Statistics ensures that the stocks are consistent over time. Figure 1.16 shows one of the main breaks in the flows between different types of contracts (TP and PT).<sup>23</sup>

Figure 1.16: Quarterly Flows: Between contract types



The transition rate between temporary to permanent was between 4% and 5% before 2005, which was consistent with the literature on contract upgrading (see Güell and Petrongolo (2007) for example). After the break it shoots up to 12% (16% following the flows calculation), almost a 200% increase. There is another spike in 2006 but it is explained by a labour market reform that happened at that time.<sup>24</sup>

<sup>23</sup>Other flow rates that suffer breaks relate to inactive workers. But since the MCVL can't speak for them, then there is nothing administrative data can add to that question.

<sup>24</sup>All temporary contracts converted to permanent before 2007 benefited from a tax exemption



The MCVL flow, on the other hand, suffer no break in 2005, but it does peak at the same time as the LFS. This could be used as supporting evidence that the TP conversion rate increased after 2005, but at a smaller rate (close to 6%). The change in the LFS must be due to the change in the survey, that implied some spurious classification errors in order to get consistent stocks.

The right panel of figure 1.16 shows starker case, where the permanent to temporary flow (PT) increases from 1% to 6% then slowly falls back to its previous level. In contrast, the MCVL only increases to 1.7% before falling after the recession. In this way the MCVL helps to interpret the LFS as coming from spurious transitions in the change in survey. The attrition problem of unemployed workers can be partly due to the same change.

## 1.6 Conclusion

Administrative datasets are a great source of information for economists, but they also present some challenges. In this chapter I analyse the case of the Spanish *Muestra Continua de Vidas Laborales* (MCVL), a rich administrative dataset of working histories of a representative sample of the Spanish workforce. This dataset is rich in information, but in its original format it has important shortcomings how it records unemployment spells. In this chapter I present a simple, systematic method to expand the original dataset by including two kinds of unemployment that the original MCVL struggles to cover: long term and short term unemployment.

Workers whose unemployment benefits expire ‘drop’ out of the sample, in that the days between benefit expiration and next employment spell are missing in the

---

scheme. Firms reacted very strongly by upgrading many temporary contracts in the last quarter of 2006. This is suggestive of firms using temporary contracts instead of permanent contracts because the former are cheaper. A simple tax rebate is enough to overcome all of the screening problems that the firm may have and would induce it to upgrade them to permanent positions.

MCVL. What I label the LTU Expansion adds these missing days. Given the rise in long term unemployment in Spain after the Great Recession, failing to include the unemployed whose benefits have expired underestimates unemployment and presents spells that are artificially shorter.

Some workers do not have the right to claim unemployment benefits. Although this is not the case of most of the unemployed, it is common among workers with short employment tenures and always the case with self-employed workers and quits. By using the information in the MCVL, I identify these cases and add the gaps between employment spells that correspond to these. I show that doing this helps match the unemployment rate from the LFS after the recession, but it is above the LFS before. I argue that given the nature of these spells the LFS would struggle to capture them, as a quarterly survey it smooths out much of this frictional unemployment. I provide further robustness checks using the LFS variable coding the relationship with the employment office and taking a closer look to the added unemployment spells.

I then use the MCVL to improve on the LFS in two main aspects: attrition bias from unemployed individuals failing to respond two consecutive quarters and changes in the survey that generate spurious labour market flows. The flows from the MCVL closely match the ones from the LFS, and where there are noticeable differences (for example in the unemployment to temporary transition rate) they help understand the sources of bias of the LFS.

Applying the proposed extensions to the MCVL makes it usable for research on unemployment. In the rest of this thesis I use this dataset for two such applications relating to long term unemployment.

# Bibliography

- Alvarez, F. E., K. Borovickova, and R. Shimer (2015). A nonparametric variance decomposition using panel data. Technical report, Mimeo, University of Chicago.
- Arranz, J. M., C. G. Serrano, et al. (2011). Are the mcvl tax data useful? ideas for mining. *Hacienda Pública Española* 199(4), 151–186.
- Bonhomme, S. and L. Hospido (2017). The cycle of earnings inequality: evidence from spanish social security data. *The Economic Journal*.
- Couch, K. A., N. A. Jolly, and D. W. Placzek (2011). Earnings losses of displaced workers and the business cycle: an analysis with administrative data. *Economics Letters* 111(1), 16–19.
- Elsby, M. W., B. Hobijn, and A. Şahin (2015). On the importance of the participation margin for labor market fluctuations. *Journal of Monetary Economics* 72, 64–82.
- García Pérez, J. I. (2008). La muestra continua de vidas laborales: una guía de uso para el análisis de transiciones. *Revista de Economía Aplicada* 16(1).
- Güell, M. and B. Petrongolo (2007). How binding are legal limits? transitions from temporary to permanent work in spain. *Labour Economics* 14(2), 153–183.
- Katz, L. F. and B. D. Meyer (1990). The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of public economics* 41(1), 45–72.

- Krueger, A. B. and A. Mueller (2011). Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data. *Brookings Papers on Economic Activity* 2011(1), 1–57.
- Laborda, A. M. (2013). La temporalidad en el mercado laboral español: nuevas aportaciones a la comprensión del fenómeno.
- Lapuerta, I. (2010). Claves para el trabajo con la muestra continua de vidas laborales.
- Moffitt, R. (1985). Unemployment insurance and the distribution of unemployment spells. *Journal of Econometrics* 28(1), 85–101.
- Silva, J. I. and J. Vázquez-Grenno (2013). The ins and outs of unemployment in a two-tier labor market. *Labour Economics* 24, 161–169.
- Sullivan, D. and T. Von Wachter (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics* 124(3), 1265–1306.
- Tattara, G. and M. Valentini (2010). Turnover and excess worker reallocation. the veneto labour market between 1982 and 1996. *Labour* 24(4), 474–500.





## Chapter 2

# Unemployment Duration Variance Decomposition à la ABS: Evidence from Spain

*Note:* This chapter is based on my work with my supervisor, Maia Güell. Maia has agreed that the essay can appear within this thesis, and that it represents a substantial contribution on my part. In particular, I carried out the data handling, main statistical work, adapted the methodology in section 2.2, provided the description in section 2.3.5, contributed to the description of the data in section 2.3 and to the analysis in section 2.4.

### 2.1 Introduction

Much has been said about the duality of the labour market in Spain in which some workers have secure permanent jobs while others have short temporary contracts (see for example Dolado, García-Serrano, and Jimeno (2002)). Another consequence of temporary contracts is the duality among unemployed workers by which the share of short term unemployment has increased but those with long durations have more

difficulties in leaving unemployment (see for instance, Güell and Hu (2006)).

In this chapter we revisit the study of the unemployment duration distribution. Existing studies of unemployment duration typically use self-reported information from labour force surveys. Instead, we use precise information on spells from administrative data. We follow the recent method proposed by Alvarez, Borovičková, and Shimer (2014) for estimating the different components of the duration of unemployment using administrative data and have applied it to Austria. In this chapter we to apply this method using Spanish Administrative data as in Chapter 1. The goal of this chapter is to compare the decomposition of the duration of unemployment of these two very different countries using comparable administrative data.

Traditionally, unemployment duration has been studied through the lense of duration models, which generally assume a parametric structure of the hazard function, distinguishing between duration dependence and unobserved heterogeneity parameters. These methods require long panel data (i.e., large  $T$ ) as well as functional form assumptions. The method proposed by Alvarez, Borovičková, and Shimer (2014) is a more statistical approach that directly estimates the relative contribution of each component to the variance of any potential duration variables, unemployment duration in particular. This approach requires fewer assumptions and as long as the data has a large number of observations (i.e., large  $N$ ), a small number of observations per individual suffice (i.e., small  $t$ ). In particular, the method can be estimated with just two completed spells of a given labour market status per individual. The nature of administrative data make this method a natural fit.

The contribution of this chapter is twofold. The first one is methodological. As explained in Chapter 1, administrative data offer several advantages with respect to standard labour force survey data. In this case, administrative data has precise information on spell duration rather than self-reported information from survey data.



But as explained in Chapter 1, the administrative data were not designed to study unemployment nor unemployment duration so they need to be augmented in order to use them for unemployment analysis. In particular, in this chapter, this affects the number of unemployed individuals observed, their unemployment duration as well as the spells being compared. As in Chapter 1, we deal with the data issues in a systematic way by using information from the Spanish Labour Force Survey as well as labour market institutions information. This is important because we find that estimates are sensitive to the treatment of these issues. The degree and nature of such data issues—which also interact with country’s labour market institutions—are country-specific. The complexity of the Spanish labour market is such that we hope that by addressing these data challenges for the Spanish case would generate a general enough guidance that is useful for applying this method to other countries. Some of these issues are dealt in Alvarez, Borovičková, and Shimer (2014) already and in this chapter we generalize this further by focusing on such complex labour market.

The second contribution of the chapter is the application to Spain, a high unemployment rate country with dual labour market dynamics, and its comparison to Austria. The Spanish case is also particularly relevant in this context because of the high quality of the Spanish social security data, featuring large  $N$  as well as large  $t$ . This also allows us to also study how the composition of the duration of unemployment changes along the business cycle. This is something that was not done in the original paper as Austria has a much more stable labour market throughout the cycle. Surprisingly we find that Spain and Austria, two very different labour markets, actually do not differ much in the differential components of the duration of unemployment.

The rest of the chapter is organized as follows. Section 2.2 explains the new method by Alvarez, Borovičková, and Shimer (2014) in detail. Section 2.3 explains

the data used and the necessary treatment done to use the data for our purposes. Section 2.4 explains our results for Spain as well as the comparison with Austria. This section also includes results along the business cycle. Finally, Section 2.5 concludes.

## 2.2 The ABS method

In this section, we explain the method proposed by Alvarez, Borovičková, and Shimer (2014) in more detail. The starting point is to assume that our outcome of interest, unemployment duration ( $y$ ), is a random variable drawn independently for each individual  $i$  from a probability distribution function  $F_i(y)$ . The population variance is then:

$$\bar{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n \int (y - \bar{\mu}_y)^2 dF_i(y) \quad (2.1)$$

Where  $\bar{\mu}_y$  is the population mean and  $n$  the population size.

The population variance can be decomposed into two components that are labelled the *within* and the *between* components. The within component is the variance of the outcome  $y$  for the average individual in the population.

$$\bar{\sigma}_w^2 = \frac{1}{n} \sum_{i=1}^n \sigma_i^2 \quad (2.2)$$

Where  $\sigma_i^2$  is the variance of individual  $i$ 's two spells. The between component represents the variance that comes from heterogeneity in the distribution functions  $F_i(y)$  across individuals.

$$\bar{\sigma}_b^2 = \frac{1}{n} \sum_{i=1}^n (\mu_i - \bar{\mu}_y)^2 \quad (2.3)$$

Where  $\mu_i$  is the mean duration of individual  $i$ 's two spells. Consider the following two extreme cases as in Alvarez, Borovičková, and Shimer (2014). First, all individuals draw from different, individual-specific distributions but there are no differences within individuals. In this case, the total variance in unemployment duration would be attributable to individual heterogeneity. Now consider the opposite case, where all individuals draw from the same distribution function, so the within variance ac-

counts for all variance and the between variance is zero as there is no heterogeneity among individuals. Imagine also that duration of unemployment does not affect the exit rate from unemployment. In this case, the total variance in unemployment duration observed would come from the distribution function  $F(y)$  itself. This component is labelled as “constant-hazard within variance” as this would be the natural model in this case.

Consider next that, while keeping all individuals homogeneous, the probability to leave unemployment depends on elapsed duration in unemployment. This means that the distribution function depends on duration  $t$ ,  $F(y, t)$ . The longer in unemployment, the longer it takes to leave unemployment. This is what the literature has referred as “negative duration dependence” and Alvarez, Borovičková, and Shimer (2014) refer to “excess within variance”.

This decomposition is done by acknowledging a key implication of the constant hazard model: the population mean and variance should be both equal to  $1/h$ , where  $h$  is the hazard rate of leaving unemployment any given period. Then we can decompose the within variance ( $\bar{\sigma}_w^2$ ) into “excess within variance” (or the difference between the population mean and variance) and “constant-hazard within variance”. The excess within variance is then:

$$\bar{\sigma}_e^2 = \frac{1}{n} \sum_{i=1}^n (\sigma_i^2 - \mu_i^2) \quad (2.4)$$

If the hazard rate is non-increasing and non-constant,  $\sigma_i > \mu_i$  and the excess variance would be positive. This is what one should expect if there is the so-called “negative duration dependence” (ie. the longer in unemployment, the longer it is going to take to find employment). The constant component is the average of the square means:

$$\bar{\sigma}_c^2 = \frac{1}{n} \sum_{i=1}^n \mu_i^2 \quad (2.5)$$

## 2.2.1 Estimation with two spells

Alvarez, Borovičková, and Shimer (2014) show that two completed spells per individual are enough to identify the three different components mentioned above. In particular, they show how to build consistent estimators for all of the above using only two spells.

Suppose that we observe two spells,  $y_{i,1}$  and  $y_{i,2}$  for individual  $i$ . Then the unbiased estimators for the individual mean and variance are

$$\hat{\mu}_i = \frac{y_{i,1} + y_{i,2}}{2}$$

$$\hat{\sigma}_i^2 = \frac{(y_{i,1} - y_{i,2})^2}{2}$$

Then the unbiased estimators for sample mean and variance are:

$$\hat{\mu}_y = \frac{1}{2n} \sum_{i=1}^n (y_{i,1} + y_{i,2}) \quad (2.6)$$

$$\hat{\sigma}_y^2 = \frac{1}{2n-1} \sum_{i=1}^n ((y_{i,1} - \hat{\mu}_y)^2 + (y_{i,2} - \hat{\mu}_y)^2) \quad (2.7)$$

As  $\hat{\sigma}_i^2$  is an unbiased estimator of  $\sigma_i^2$ , then the unbiased estimator of the within variance ( $\bar{\sigma}_w^2$ ) is

$$\bar{\sigma}_w^2 = \frac{1}{n} \sum_{i=1}^n \hat{\sigma}_i^2 \quad (2.8)$$

For the between variance component we need unbiased estimators for the squares of the individual al populations means ( $\mu_i^2$  and  $\bar{\mu}_y^2$ ) as shown in equation 2.3. These can be shown<sup>1</sup> to be:

$$\mathbb{E}\mu_i = \hat{\mu}_i^2 - \hat{\sigma}_i^2$$

$$\mathbb{E}\bar{\mu}_y = \hat{\mu}_y - \frac{1}{2n} \hat{\sigma}_y^2$$

---

<sup>1</sup>For details refer to Alvarez et al. (2014) section 3.2.

So the unbiased estimator of the between component is:

$$\bar{\sigma}_b^2 = \frac{1}{n} \sum_{i=1}^n (\hat{\mu}_i^2 - \hat{\sigma}_i^2) - (\hat{\mu}_y - \frac{1}{2n} \hat{\sigma}_y^2) \quad (2.9)$$

Similarly, we can reformulate equations 2.11 and 2.10 to obtain:

$$\hat{\sigma}_e^2 = \frac{1}{n} \sum_{i=1}^n (\hat{\mu}_i^2 - \frac{1}{2} \hat{\sigma}_i^2) \quad (2.10)$$

$$\hat{\sigma}_c^2 = \frac{1}{n} \sum_{i=1}^n (\frac{3}{2} \hat{\sigma}_i^2 - \hat{\mu}_i^2) \quad (2.11)$$

In the particular case where we decompose the natural logarithm of the duration of unemployment, the constant hazard component is always equal to  $\pi^2/6$ , while the excess variance estimator will then be

$$\hat{\sigma}_e^2 = \hat{\sigma}_w^2 - \hat{\sigma}_c^2 = \frac{1}{n} \sum_{i=1}^n \hat{\sigma}_i^2 - \frac{\pi^2}{6} \quad (2.12)$$

For all of this analysis it is crucial to have a large number of observations (large  $n$ ) so that the estimators are as close as possible to the true components. Notice how the number of individual observations  $J$  is not important as long as the  $n$  is large. In this way a long panel is not necessary for the estimation.

More details can be found in Alvarez, Borovičková, and Shimer (2014).<sup>2</sup>

## 2.3 Data

As in Chapter 1, we use the Spanish Social Security data. It is called the “Spanish Continuous Working Life Sample” (MCVL, the Spanish acronym, hereafter) which comprises the working histories of a 4% sample of the working population for the

---

<sup>2</sup>In particular, they relax the assumption that the distribution of the two spells is the time-invariant. In this case, regressing duration on the number of spell is enough to control for changes in  $F_i$  through time.

years 2004-2013. Relevant to this chapter is the large number of observations both in terms of individuals at any point in time as well as the large number of years available. As explained in Chapter 1 these data were not designed to study the labour market, they were designed to have an account of employed workers (who contribute to the social security) and of unemployed workers entitled to unemployment insurance (who have to be paid by the social security). Therefore the data will display blanks whenever workers are not in one of these two categories. The data needs to be augmented before using to study the labour market. In the following subsections, we explain again the key data expansions explored in Chapter 1 needed for this particular application explored in this chapter. As in Alvarez, Borovičková, and Shimer (2014), for this application we will need the first two consecutive spells of unemployment observed in the data for all individuals.

### **2.3.1 Data expansions**

As in Chapter 1 (Lafuente (2017)), we will make data expansions in a systematic way by using information from the Spanish Labour Force Survey as well as labour market institutions information. In particular, we will do the following 3 expansions to deal with: unemployment duration, the number of unemployed workers observed and the spells being compared.

### **2.3.2 Duration of unemployment in Administrative data**

The first issue of administrative data is with respect to unemployment duration. Recall that only unemployed workers receiving unemployment insurance are observed in the data. Once their benefits run out, if they are still unemployed, they appear as blank in the data. This implies that unless the data are augmented, unemployment duration will be underestimated. Table 2.1 illustrates this point with an example of a worker during three consecutive spells. The first row displays the labour market status of a worker in reality and the second row displays what appears in the administrative data.

The data displays cases that look as the example in Table 2.1. There is a worker with a first spell in unemployment and receiving unemployment insurance. In this case both the data and reality coincide. In the second spell, the data are *blank* while the reality could be that the worker is still unemployed but does not receive unemployment insurance anymore or that the person went out of the labour force. Finally, in the third spell the worker is employed and the data reflects this.

Table 2.1: Duration of unemployment

Example. LTU expansion

	spell 1	spell 2	spell 3
Reality:	U with UI	(i) U run out UI, (ii) out LF	E
Data:	Registered U	<i>blank</i>	E

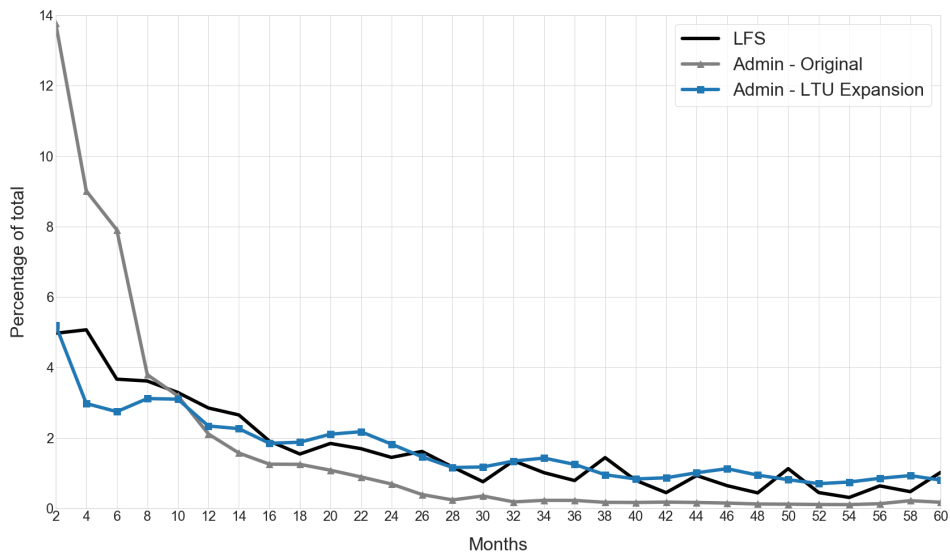
Notes: U stands for unemployed, UI for unemployment insurance, LF for labour force and E for employed.

The key here is spell 2. If indeed the worker continued to be unemployed (without insurance), then unless the data are augmented we would be underestimating unemployment duration (only counting spell 1). It is likely that in this example the worker indeed continued to be unemployed given that in spell 3 the worker is back to employment. Also, given the fact that we concentrate on the first two completed spells of unemployment observed in the data for any individual this implies that generally workers are going to be relatively young and inactivity is a less likely event. As stated, we will look for information in the labour force survey in order to fill this blank in the data. Using the Spanish LFS we find that as many as 88% unemployed workers who run out of unemployment insurance continue to be unemployed in the period afterward. Accordingly, as a first approximation we replace all the blanks of cases as in Table 2.1 with unemployment status. Admittedly, this is a broad brush

solution but it turns out to be very successful in this case. As explained in Chapter 1, this approach was also used by García Pérez (2008).

After implementing this data expansion that we label as LTU expansion, we document the following three facts. First, the unemployment duration data in the LFS and the MCVL become more comparable. Figure 2.1 displays a histogram of the duration of (incompleted) unemployment in months for 2013 (third quarter) with the LFS data, the original MCVL and the MCVL after the LTU expansion. As can be seen the original MCVL puts more weight into shorter durations than the LFS, but after the addition, the LFS and the MCVL are much more comparable.<sup>3</sup>

Figure 2.1: Duration in months of (uncompleted) unemployment in 2013 (III), LTU expansion

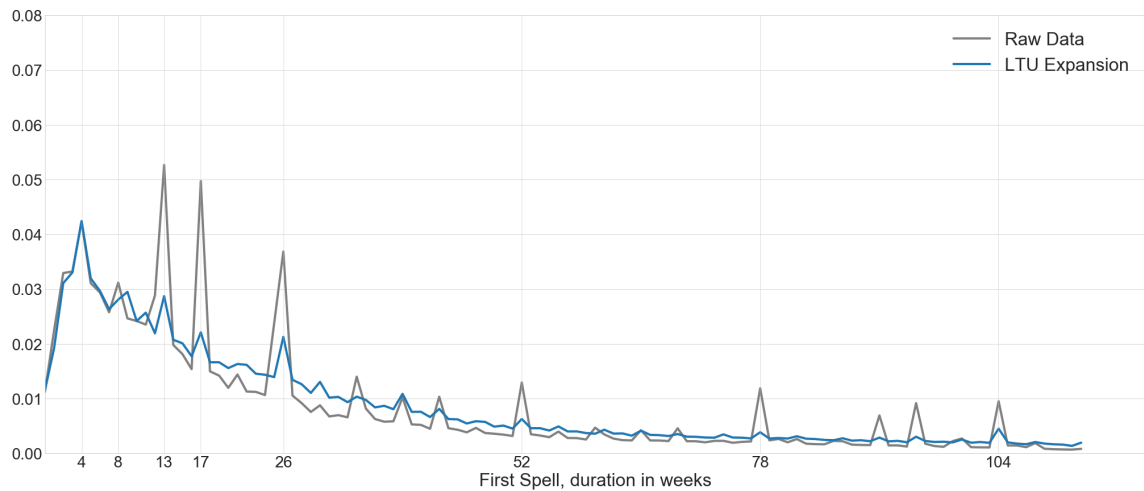


<sup>3</sup>Note however that the duration reported in the LFS has issues too: it is self-reported, the data frequency is quarterly and there are problems with attrition for unemployed workers (see section 1.5.1 in Chapter 1).



Second, we focus on the duration of the first completed unemployment spell observed for each individual in the data. As Figure 2.2 shows, the original MCVL displays pronounced spikes that coincide with the end of unemployment benefits (e.g., at 3, 4 and 6 months), see Chapter 1. This again, implies underestimating duration as workers may continue unemployed without benefits. The figure also shows that after the LTU expansion, these spikes are smaller.

Figure 2.2: Duration in weeks of completed unemployment spells, 2005-2013, LTU expansion



Third, as shown in Chapter 1, the unemployment rate series using the MCVL after this addition becomes closer to the LFS (see Figure 1.4 in Chapter 1).

Other cases that appear in the data are such that there is a first spell of registered unemployment (i.e., worker is unemployed and receives unemployment insurance) followed by a blank thereafter. That is, the first and second spell look like the example in Table 2.1. However, note that these are cases displaying right-censoring. Also note that for this application, we will not be using such observations as the minimum requirement is two *completed* spells.

### 2.3.3 In between employment spells

We now turn to the second issue of administrative data, which relates to the labour force status in between two observed employment spells. This is likely to be unemployment spells without unemployment insurance. It could also be job-to-job movements. To the extent that this is an unemployment spell, unless the data are augmented, the number of unemployed workers will be underestimated. Table 2.2 illustrates this point with an example of a worker during five consecutive spells (in order to include two potential completed spells of unemployment). As before, the first row displays the labour markets status of a worker in reality while the second row displays what appears in the administrative data.

Table 2.2: Number of unemployed

Example. STU expansion

	spell 1	spell 2	spell 3	spell 4	spell 5
Reality:	E	U no UI, JtoJ or out LF	E	U no UI, JtoJ or out LF	E
Data:	E	<i>blank</i>	E	<i>blank</i>	E

Notes: U stands for unemployed, UI for unemployment insurance, JtoJ for job-to-job transitions and E for employed.

The data displays cases that are two employment spells with a blank in between, as in the example in Table 2.2. There is a worker with several spells in employment, in which case the data and reality coincide. These employment spells are in-between other spells in which the data are blank. These spells could be unemployment without insurance or they could also be job-to-job transitions or even spells out of the labour force. It is crucial to be able to distinguish between these possibilities as much as possible, otherwise the number of unemployment spells could be underestimated.

We proceed as follows. We can distinguish among two cases. Please note that the MCVL provides two key pieces of information: the duration of the spell and the identity of the firm when employed. The first case corresponds to employment in spells 1 and 3 (or spells 3 and 5) in the same firm. In this case, as explained in Chapter 1, we would consider this as a recall and, therefore, spell 2 (or spell 4) would not be considered as unemployment nor out of the labour force.

The second case corresponds to employment in spells 1 and 3 (or spells 3 and 5) in different firms, and the spells in between employment spells (i.e. spells 2 or 4). In this case, it could be unemployment, (longer) job-to-job transitions or out of the labour force. One solution to these type of observations would be to treat them as *non-employment* as in Alvarez, Borovičková, and Shimer (2014). In this chapter, however, we would like to go further in trying to distinguish unemployment from other situations by using complementary information. In this case, we use institutional information, that is, legislation on unemployment insurance that establishes who is eligible for unemployment insurance or not. We note that in this case, these blank spells in the data are fairly short and in this sense information in the labour force survey is not that useful since the Spanish LFS is quarterly and short spells are not easy to identify. However, legislation on unemployment insurance allows us to assign individually whether these spells correspond to situations with no right to unemployment insurance. Those that do have the right of unemployment insurance, we consider them as either job-to-job transitions or out of the labour force. The reason being that if they were unemployed the administrative data would capture that since they have benefits. However, for those that do not have the right of unemployment insurance it could be one of the following three cases: (i) wage earners with too short accumulated tenure in past employment (usually coming from temporary contracts); (ii) self-employed workers ; or (iii) workers that have quit. In these cases, we replace the blanks at the individual level with unemployment status. We find that around 81% of the blanks in this category do not have the right of unemployment insurance

and thus we replace the blanks with unemployment status. While 19% of the blanks in this category do have the right of unemployment insurance and are therefore not replaced by unemployment status.

In sum, this data expansion implies that the blanks in Table 2.2 could be either be a recall, a job-to-job movement, out of the labour force, or an unemployment spell. In the first three cases, the data is left unaffected. But in the latter we replace the blank with an unemployment spell. After implementing this data expansion that we label as *short term unemployment addition* (STU expansion), we document the following three facts:

1. First, we add the new data to the previous Figure 2.1 and get a reduce the discrepancy with respect to the LFS, especially for short durations, see Figure 2.3 below.
2. Second, we add the new data to the previous Figure 2.2 and the pronounced spikes that coincide with the end of unemployment benefits get further reduced, see Figure 2.4 below.
3. Third, as shown in Chapter 1, the unemployment rate series using the MCVL after this addition become closer to the LFS (see Figure 1.7 in Chapter 1).

### **2.3.4 On the spell number**

In order to implement the method proposed Alvarez, Borovičková, and Shimer (2014), we select the first two completed unemployment spells observed for every individual. This is in principle straightforward, except that it can be the case that some unemployment spells are not detected in the original data as explained in the previous subsection. Once the data have been augmented as explained in subsection 2.3.3, it can be the case that the first two spells observed are not the same as the first two spells in the original raw data.

Figure 2.3: Duration in months of (uncompleted) unemployment in 2013 (III), STU expansion

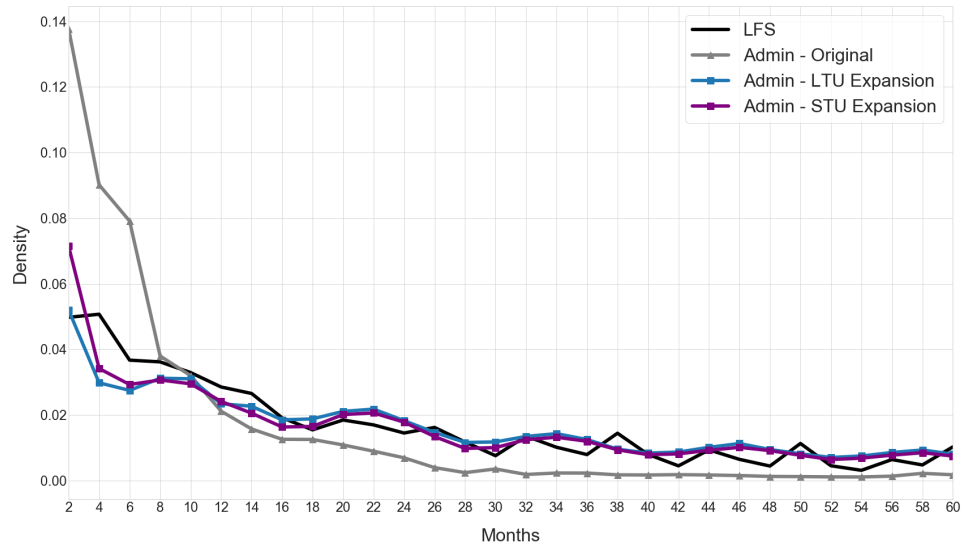


Figure 2.4: Duration in weeks of completed unemployment spells, 2005-2013, STU expansion

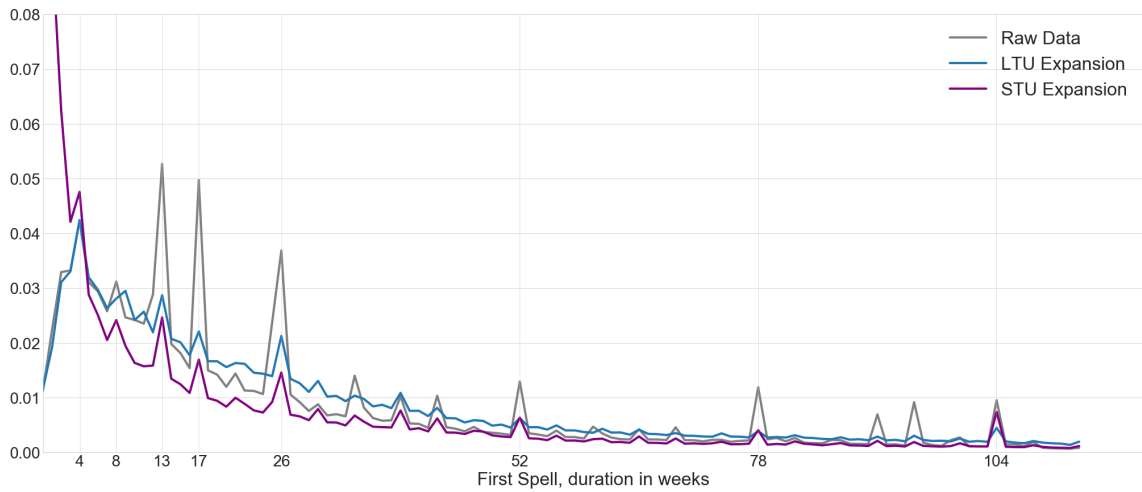


Table 2.3 illustrates this point with an example of a worker during six consecutive spells (in order to include different first two spells of unemployment). The first row display the labour market status of a worker in reality. This worker has a spell of employment followed by a spell of unemployment without benefits and then followed by two sets of employment-unemployment (with benefits) spells. The second row displays the labour market status of a worker according to the administrative data. The data and reality coincide except for the second spell which is blank in the data.

Table 2.3: Spell Number Correction

Example. STU expansion

	spell 1	spell 2	spell 3	spell 4	spell 5	spell 6
Reality:	E	U no UI	E	U w/UI	E	U w/UI
Data:	E	<i>blank</i>	E	Registered U	E	Registered U
Raw Data:	E	<b>“E”</b>	E	Registered U <sub>(spell #1)</sub>	E	Registered U <sub>(spell #2)</sub>
STU expansion:	E	U <sub>(spell #1)</sub>	E	Registered U <sub>(spell #2)</sub>	E	Registered U

Notes: U stands for unemployed, UI for unemployment insurance, JtoJ for job-to-job transitions and E for employed.

The third row in the table shows what would happen if we were to work with the raw data, that is if the data was *not* augmented. The first two spells of unemployment that would be identified would be spells number 4 and 5. This means that implicitly that we would be assuming that the second spell is an employment spell (thus labeling it “E”). Instead, as shown in the fourth row of this table, if we augment the data explained in subsection 2.3.3, we would be replacing the blank with an unemployment spell (note in this case we have assumed this is the reality as per the first row). This would imply that the first two spells of unemployment would now be spells 2 and 4 rather than spells 4 and 6.

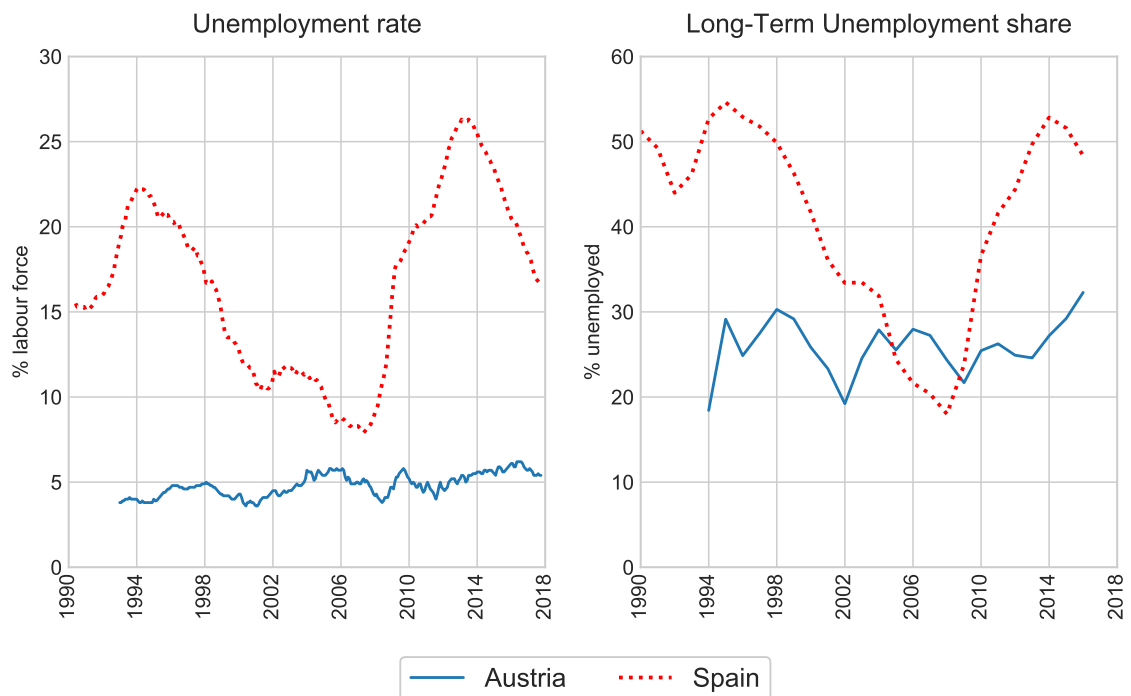
This example has meaningful economic content. The case described corresponds to employment spells of quite short duration with no right of unemployment insur-

ance. In Spain these correspond to spells out of temporary contracts which are very predominant (35% of the workforce is on temporary contracts). This means that we can somehow approximate what is the effect of temporary contracts in Spain to the question at hand. Using the raw data without the STU expansion approximates the lack of temporary contracts in the economy to some degree and indeed employment would appear as more stable than it actually is (i.e., some unemployment spells would implicitly be considered as employment in the context of this chapter).

### 2.3.5 Austria and Spain

In the results section we will be comparing Austria and Spain. To this end, it is worth drawing attention to the key labour market differences between these two countries.

Figure 2.5: Unemployment and Long Term Unemployment, Austria and Spain



Notes: Data from OECD (2017)

Figure 2.5 shows the unemployment and long-term unemployment<sup>4</sup> rates for both countries since 1990 (1994 in the case of Austria). The first thing to note is that Spain has substantially higher unemployment than Austria, and its volatility is higher as well. While for Austria unemployment has stayed between 4 and 5%, in Spain unemployment fluctuates between 8% in 2007 and 26% in 2013. The right panel also shows that the share of long-term unemployment is more volatile in Spain, going from over 50% in 1990 to under 20% in 2007. Long-term unemployment is more relevant in Spain, but in the 1995-2008 period it decreased considerably, partly due to the widespread adoption of temporary contracts (see Güell and Hu (2006)). Austria's long term unemployment is one of the lowest of the European Union, with similar levels to Denmark, Luxembourg and Sweden. It's share over total unemployment is also fluctuating between 20% and 30%. Business cycles are clearly marked in Spain, while in Austria fluctuations are small and close to a trend.

Unemployment insurance is also different in both countries. Table 2.4 compares minimum requirements to claim (in the form of previous contribution to the social security), generosity as measured by the replacement rate (percentage of previous wages) and maximum duration of entitlement. Overall, Austria has a less generous system, with a constant replacement rate of 55% of average net wages over the previous year, while Spain has very generous short-term benefits (70% over the average net wage over the last 3 months), but lower after 3 months. The main difference comes with time of entitlement duration: in Spain is proportional to contribution period (a month for every 3 months of employment) and up to two years, while in Austria, for those younger than 40 years of age it's only 6 months. Workers older than 40 and 50 years of age with enough contribution periods can be granted up to a year.

In both countries workers have to report to the employment office and regularly

---

<sup>4</sup>The OECD defines long-term unemployment as 'people who have been unemployed for 12 months or more'.



Table 2.4: Structure of Unemployment Insurance

	Austria	Spain
Minimum contribution period	12 months in 2 years <sup>5</sup>	12 months in 6 years
Replacement rate	55%	70% in the first 3 months, 50% after
Maximum duration	1 year <sup>6</sup>	2 years <sup>7</sup>

Notes: Data from the European Commission

meet an employment advisor to check they are actively looking for employment and show that they have not declined any suitable job offer. In the case of expiration of unemployment insurance, both countries offer unemployment assistance benefits for those who run out of unemployment insurance. In Austria, these can last up to a year or indefinitely if certain conditions are met. In Spain, it can last up to 11 months or up to 18 months if the individual has family responsibilities or if she's older than 55. In both cases the amount of benefits is a share of the minimum living income calculated by both countries and independent of previous earnings.

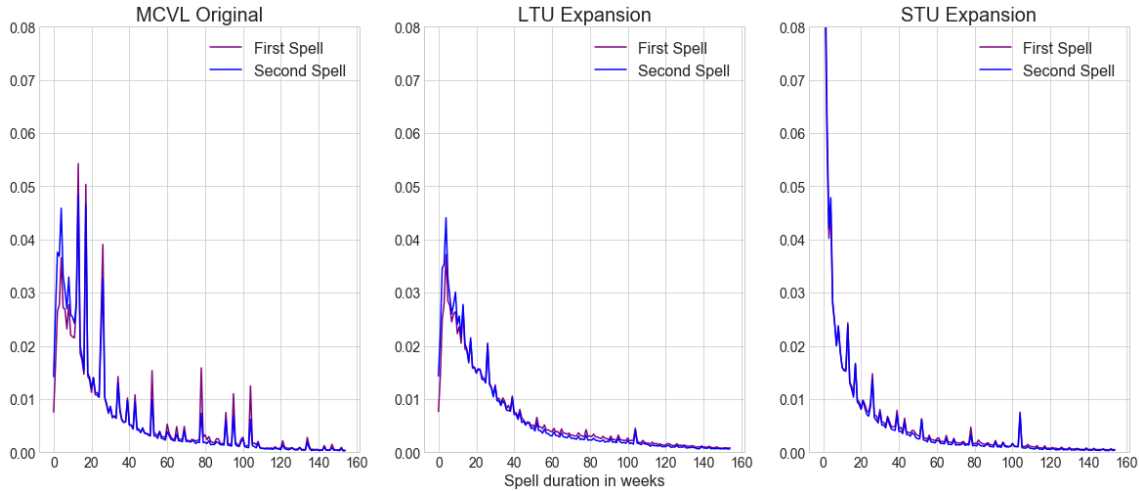
## 2.4 Results

In this section we explain our results. Before that, we need to check a necessary requirement to implement the Alvarez, Borovičková, and Shimer (2014) method. As explained, the method uses the first two spells of unemployment observed for each individual. The requirement is that these two spells are not too different. This means that the distribution from which the first spells are drawn is not too different from the distribution from which the second spells are drawn. We check this empirically in our data as in Alvarez, Borovičková, and Shimer (2014).

Figure 2.1 displays the hazard rate of the first spells (blue line) as well as of the

second spells (purple line).

Figure 2.1: Hazard Rate By Number Of Spell



The left hand side figure displays the hazard rates of the first and second spells using the raw data. As it can be seen, the two spells are very similar. As discussed before, there are some pronounced spikes that correspond to unemployment benefit expiration. The middle and right hand side figures display again the hazard rates of the first and second spells using the data that has been expanded with the LTU expansion (middle) and the data that has been expanded with the LTU and STU expansions explained in the previous section. We find the same, that is, the two spells are very similar. It can be seen in the last two graphs that the expansions of the data make the spikes smoother.

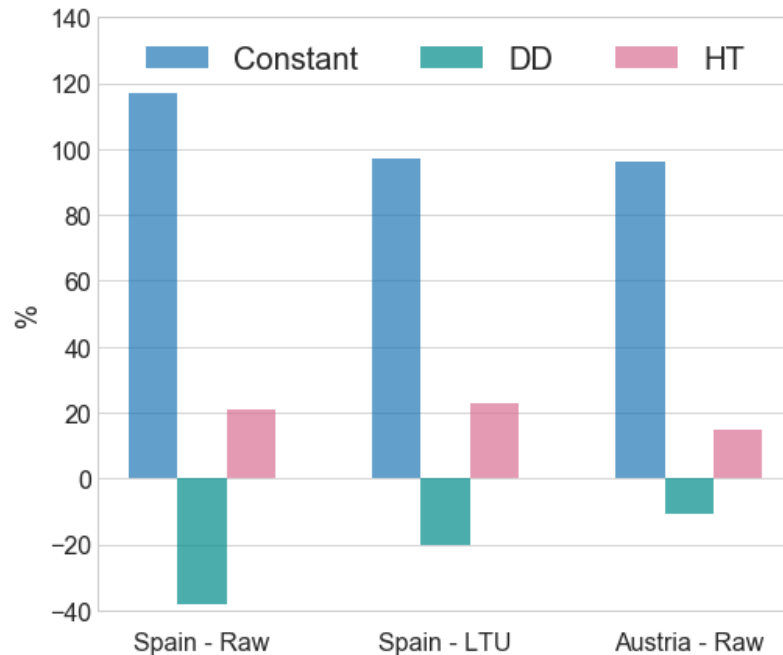
In the next subsections, we proceed to report the variance decomposition of unemployment duration (in logs) with the raw data as well as with the different expansions of the data in order to highlight the importance of such data expansions. As will be seen, results are sensitive to such expansions so it is crucial to undertake them. As explained in section 2.2 we will distinguish between three different components of the variance of unemployment durations across individuals.

These three components are: heterogeneity (variation across individuals), duration dependence (variation within individuals due to different elapsed durations) and the “constant” component consistent with homogeneous workers and no duration effects.

### 2.4.1 Variance decomposition with raw data and LTU expansion

We start decomposing the variance of unemployment duration with the raw Spanish data. Figure 2.2 displays the results in the left hand side figure. For comparison we also display the equivalent results for the raw Austrian data in the right hand side figure. The detailed figures can be found in Table B.1 in Appendix B.

Figure 2.2: Decomposition with LTU addition



Notes: Unemployment duration in logs. Results for Austria taken from Alvarez, Borovičková, and Shimer (2014). DD stands for duration dependence, HT stands for heterogeneity.

There are several aspects to highlight. First, the broad decomposition in Spain

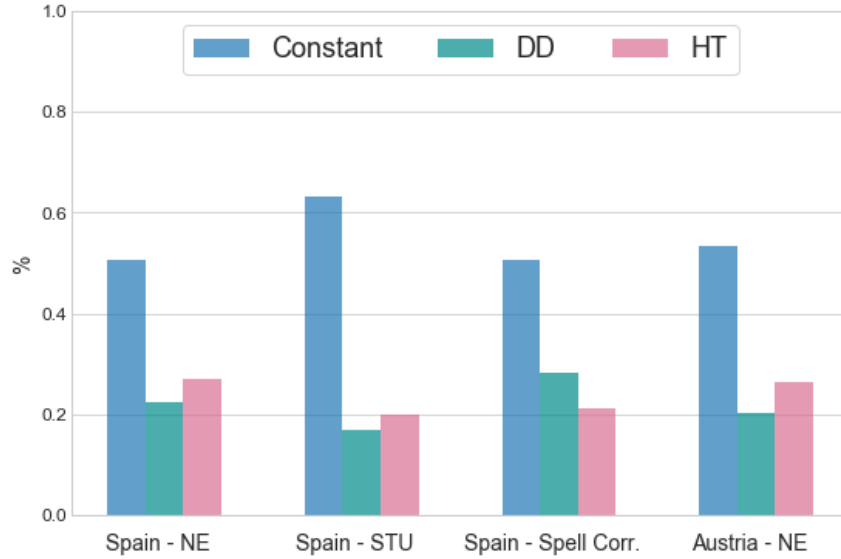
and in Austria is surprisingly similar. For both countries, the component that stands out is the “constant” component. Secondly, for both countries the “duration dependence” component is actually a negative number. This is an artifact of the untreated data. That is, the raw data display unemployment spells that are too short as they only capture unemployment while receiving benefits. This underestimation of unemployment duration appears as if exit sharply increases around benefit expiration. This effect is strong enough to generate a positive effect on exit and is present both in Spain and Austria, but is more pronounced in Spain due to the different functioning of the two labour markets as explained in section 2.3.5. The longer unemployment benefits and longer unemployment durations as explained in Spain compared to Austria can reconcile this finding. As explained above this shortcoming of the administrative data can be addressed with the LTU expansion. Indeed, the middle figure shows the decomposition once the data have been augmented in this way. As can be seen, the anomaly of the “duration dependence” component is reduced (becoming a higher number, even if still a negative number). And the shares of the within component get redistributed, while as one would expect, the heterogeneity component remains unaffected. Finally, the share of the variance explained by heterogeneity is rather small compared to the constant component. And, again, this is a common feature in both countries.

## 2.4.2 Variance decomposition with all expansions

We now decompose the variance of unemployment duration with the Spanish data fully expanded. We also add the case where we count all non-employment spells as ABS do for Austria. Figure 2.3 displays the results. More detailed results can be found in Table B.2 in Appendix B. The first left hand side figure presents the decomposition for Spain counting all non-employment gaps (*Spain - NE*). The second figure presents the results for the the LTU and STU expansions (*Spain - STU*). The third figure also takes into account the correct spell number that follows the

STU expansion as explained in section 2.3.4 (*'Spain - Spell Corr.'*). The last figure presents the results for Non-Employment spells in Austria.

Figure 2.3: Decomposition with LTU and STU additions



Notes: Unemployment duration in logs. Results for Austria taken from Alvarez, Borovičková, and Shimer (2014). DD stands for duration dependence, HT stands for heterogeneity.

Figures 1 and 4 are roughly similar. They are both however very different to Figure 2.3 in the previous section. The main difference is that the duration dependence component is now a positive number, as one would expect. The STU expansion has made the difference in this respect. This is because the raw data display a hump-shape at very short durations. That is, exit from unemployment is increasing with duration for very short durations but this disappears when the data have all expansions (see Figure 2.1). The STU addition incorporates more short spells into the data, and so does the Non-Employment case as it is a less restricted STU correction. Still, the main result is that the main component is the constant component, while heterogeneity and duration dependence are relatively similar. Again, the two countries are very similar in terms of the decomposition of non-employment dura-

tion. By adding only the cases of workers who have no right to claim unemployment benefits, the STU expansion allows us to refer to it as unemployment rather than non-employment as in the case of Austria. The STU addition is different from all Non-Employment in that the constant component gains relevance at the expense of both duration dependence and heterogeneity (each losing about 5 percentage points). However the ranking and size of heterogeneity and duration dependence is the same. This is not true when considering the Spell Correction.

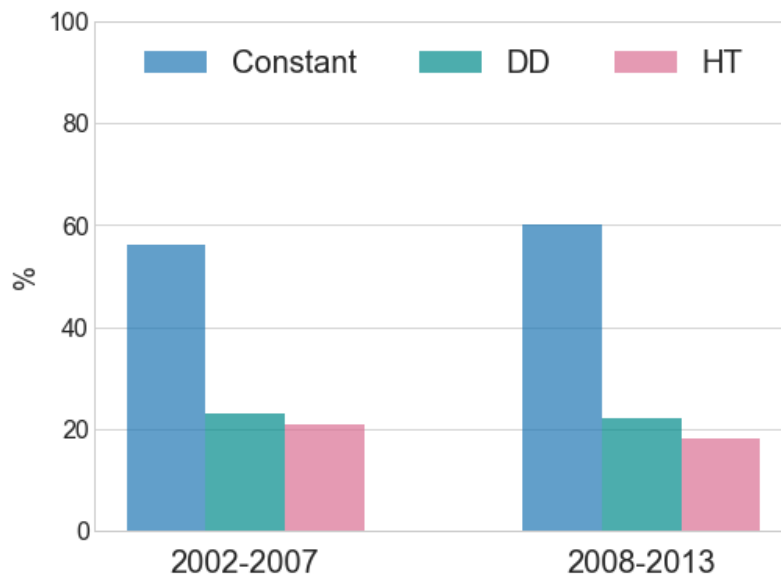
This is a very interesting result. As explained before, if we do not correct for the spell number it is as if we were somehow suppressing the impact of temporary contracts in the Spanish labour market. Note that in the STU expansion (second and fourth figures), Spain and Austria have the same ranking (constant > heterogeneity > duration dependence), with duration dependence and heterogeneity being roughly similar. Instead, once we correct for the spell number (i.e. and the impact of temporary contracts is present through short unemployment spells), then Spain and Austria become less similar, and duration dependence in Spain gains relevance. This is consistent with Güell and Hu (2006) that find that temporary contracts in Spain increase duration dependence.

### **2.4.3 Business cycle comparisons**

An important aspect of the Spanish data and labour market is that we can compare unemployment between different moments of the business cycle, which display large variation in unemployment rates as well as unemployment durations. The question is, is the decomposition of the unemployment duration very different along the business cycle? Figure 2.4 displays the decomposition for years 2002-2007 and 2008-2013 (left hand side and right hand side figures, respectively). The former years were years of lower unemployment and share of long-term unemployment (around 10% and 30%, respectively), while the latter are years of higher unemployment (around

25% and 40%, respectively).

Figure 2.4: Spain along business cycle



Notes: Unemployment duration in logs, Spain, fully expanded data. DD stands for duration dependence, HT stands for heterogeneity.

As can be seen the importance of the three different components is surprisingly similar in the two very different periods considered. Table B.3 in Appendix B displays the numerical results. As before, the constant component is clearly the most important, followed in similar weights by the heterogeneity and duration dependence components. It is worth mentioning that here we have not distinguish between observed and unobserved heterogeneity as traditional models estimating duration models do (see for example Lancaster (1979) and Bover, Arellano, and Bentolila (2002)). It is possible that if we did that, the importance of unobserved heterogeneity would be different along the business cycle. But it is likely that the comparison with respect to the constant and the duration dependence components remains unchanged and similar along the business cycle.

## 2.5 Conclusions

In this chapter, we have applied the method recently developed by Alvarez, Borovičková, and Shimer (2014) to Spain using the administrative social security data described in Chapter 1. The contribution of this chapter is methodological as well as being economically relevant.

Methodologically, as explained, administrative data have several advantages, namely, in this case, precise spell information as opposed to self-reported duration from survey data. However, administrative data need to be expanded in order to use them for research. The different expansions needed are country-specific. The complexity of the Spanish labour market is such that both short and long unemployment durations coexist and we hope that by addressing these data challenges for the Spanish case this will provide general enough guidance to be useful for applying this method to other countries. Specifically, workers who remain in unemployment after their benefits expire will have durations that are too short in the administrative data. At the other extreme, workers without benefit entitlement who experience short unemployment durations will be lacking a recorded unemployment spell. This is relevant because, as explained, the results of the method proposed by ABS are sensitive to the data expansions.

From an economic point of view, we have compared two very different countries (Austria and Spain) as well as Spain along the business cycle. The different data expansions are also related to underlying functioning of the labour market. In particular, we have seen how the STU expansion relates to the presence of temporary contracts in Spain. This allows us to approximate how these contracts may relate to the decomposition of unemployment duration. In a nutshell, we find that the decomposition of unemployment duration is very similar both between countries (especially when minimizing the incidence of temporary contracts) as well as along the



business cycle within Spain. This may be an unintuitive result at first. But upon reflection these results could be rationalized by a model of the labour market in which the overall level of unemployment and unemployment duration do not affect the importance of the different components of unemployment duration. We leave this question for future research.



# Bibliography

- Alvarez, F., K. Borovičková, and R. Shimer (2014). Decomposing duration dependence in the job finding rate in a stopping time model. Mimeo University of Chicago.
- Bover, O., M. Arellano, and S. Bentolila (2002). Unemployment duration, benefit duration and the business cycle. *The Economic Journal* 112(479), 223–265.
- Dolado, J. J., C. García-Serrano, and J. F. Jimeno (2002). Drawing lessons from the boom of temporary jobs in Spain. *The Economic Journal* 112(480).
- García Pérez, J. I. (2008). La muestra continua de vidas laborales: una guía de uso para el análisis de transiciones. *Revista de Economía Aplicada* 1(16).
- Güell, M. and L. Hu (2006). Estimating the probability of leaving unemployment using uncompleted spells from (repeated) cross-section data. *Journal of Econometrics* 1(133).
- Lafuente, C. (2017). The best of two worlds: Assessing the use of administrative data for the study of unemployment using the labour force survey as a benchmark. Mimeo University of Edinburgh.
- Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica: Journal of the Econometric Society*, 939–956.
- OECD (2017). Long-term unemployment rate (indicator).



# Chapter 3

## Search capital and Unemployment Duration

### 3.1 Introduction

Most classical explanations of long term unemployment (LTU thereafter) relate mostly to older workers, whether is by depreciation of their human capital due to an exogenous shock, as in Ljungqvist and Sargent (2008), or because they are better insured (unemployment benefits, own savings) against unemployment. However, these explanations don't apply so well to younger workers, which have low tenures (and thus have lower benefits) and have not yet accumulated much human capital. In the last recession the young have been hit harder by long term unemployment, as figures 3.1 and 3.2 shows for different European countries. This chapter introduces a novel mechanism that can help explain these patterns by treating job finding as a skill that is learned and forgotten over time. This skill is separate from traditional human capital because it is unrelated to workers productivity on the job, but it accumulates process. I refer to this skill as search capital.

In particular, I characterise this skill as the one that make workers searching for employment more successful by finding jobs faster. Workers with high search capital receive more offers in any given period and thus have a wider choice of jobs, making them more likely to find a better job, faster. Because of this they will be less likely to remain unemployed for long periods of time. Workers' search capital increases by successfully finding a job, in the sense that workers use their previous experience (the places they applied to, the way they pass the different stages of recruitment processes) to search more efficiently next time they face unemployment. While some of these skills can be learnt after a failed application workers are likely to learn more through successful search. However search capital deteriorates if it is not used, as recent experiences are more relevant than those in the distant past. This implies that proficient searchers are the ones that have had more recent exposure to search, while new entrants to the job market and workers with long tenures are going to be relatively worse at finding jobs.

In this chapter I will be focusing on the case of Spain, the country in figures 3.1 and 3.2 where the increase in youth LTU has been more pronounced after Greece. The dual labour market structure that characterises Spain generates substantial heterogeneity in labour market experiences among the unemployed, which can be used to identify search capital. Workers in temporary contracts, which rarely ends in promotion to regular contracts, are forced to constantly search for employment. Workers with permanent contracts enjoy longer job tenures and rarely experience unemployment and thus do not accumulate search capital. Using administrative data for Spain, I find a negative correlation between the number of temporary jobs and unemployment duration, which I use as proxy for search capital. I also test whether workers with more temporary contracts in the past find worse jobs in the future by

Figure 3.1: Youth and overall long term unemployment

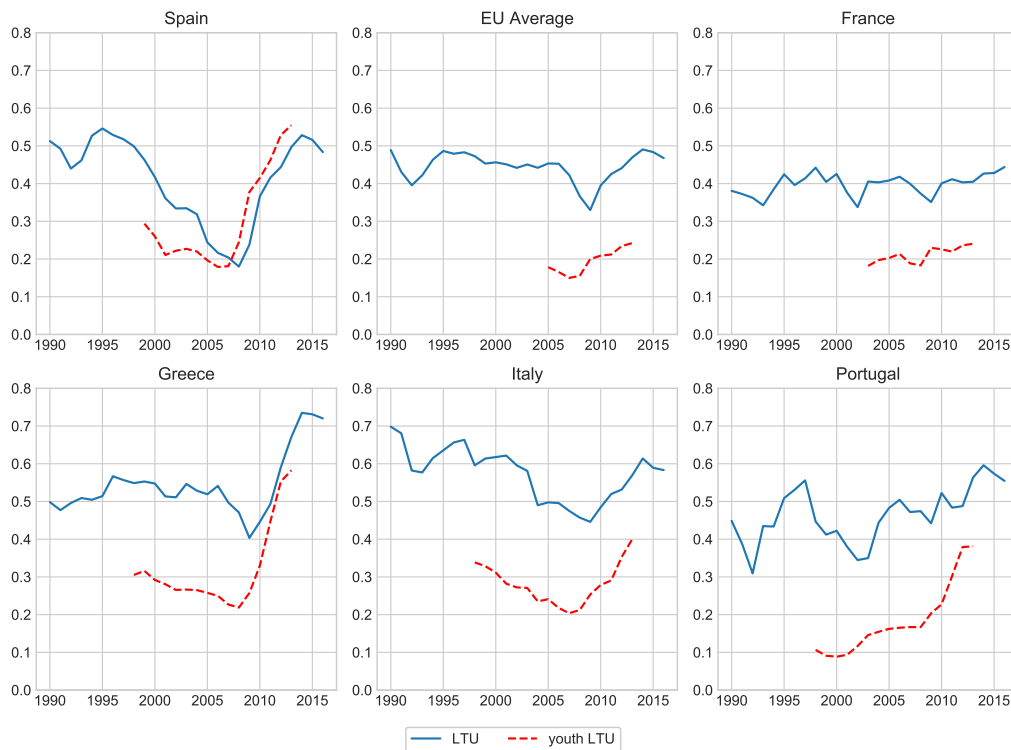
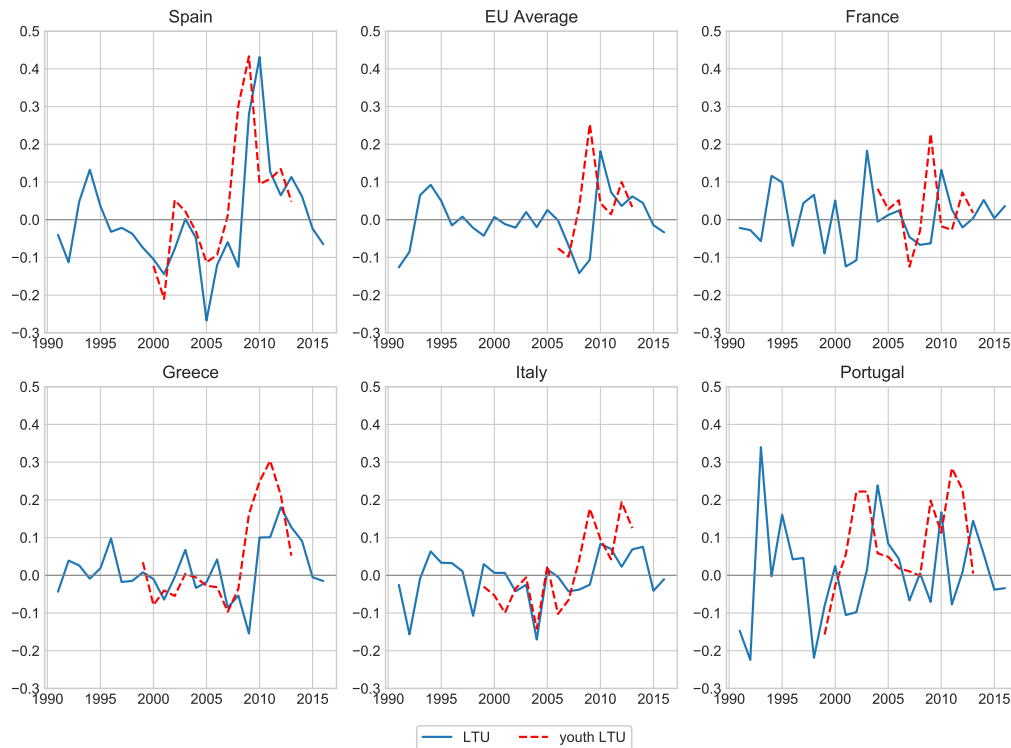


Figure 3.2: Youth and overall long term unemployment, annual log change



Source: OECD (2017)

looking at wages and duration of the next job. I find a small but positive effect on wages, and a positive effect on duration after controlling for individual fixed effects. This provides some evidence at the individual level to support the search capital channel.

I then propose a dynamic search model to quantify the effect that search capital differences have on aggregate outcomes. In particular, I focus on the differences in job finding rates and unemployment duration among different age groups. The model generates life-cycle dynamics that closely match the data. The addition of search capital helps explain the labour market flows of young workers and generates substantial differences in unemployment durations among workers of the same age group. Over time, workers become worse at searching as they settle into stable jobs, their search skills deteriorating as a result. This could potentially pose a problem if these workers were to lose their jobs, as an inflow of inefficient searchers can result in more long term unemployment.

The rest of the chapter is structured as follows: Section 3.2 explains search capital in more depth and compares it to other mechanisms in the literature that may have similar effects; Section 3.3 provides empirical evidence at the individual level; Section 3.4 presents a theoretical search model that incorporates search capital and calibrates it for Spain; Section 3.5 concludes.

## **3.2 Search Capital and Long Term Unemployment**

This section explains more extensively what I refer to as search capital and how it relates to long term unemployment (LTU thereafter). For this the interactions with a dual labour market are very relevant, so this section also argues that the



case of Spain is relevant to test the effects of search capital. It then identifies and reviews the main competing mechanisms that could drive an increase in long term unemployment and have similar effects to those that would be explained by search capital.

### **3.2.1 Search Capital, Dual Labour Markets and LTU**

Earlier I defined search capital as the set of skills that help workers find jobs. For example, knowing the places they should applied to (applying for the right kind of jobs for the productive skills workers have, diversifying their search, etc) or knowing how to prepare for the different stages of recruitment processes (interviews, tests, etc). While some of these skills can be learnt after a failed application (a disastrous interview can help improve next one) workers are likely to learn more through successful search - which means they can use their previous experience should they need to find a job again. Being a more efficient searcher translates into being able to apply to more jobs and increases the chances of being offered the job after the application process. That is the approach I follow in the theoretical model of Section 3.4.

This treatment has some advantages: first, it makes a mapping between an unobservable variable (search capital) and an observable outcome: number of successful job searches or jobs held by the worker. Second, it makes search capital dynamics easier to incorporate to model, while modelling as a learning process through fail applications as well can become more complicated and potentially imply that search skills increase with time in unemployment. It is a well known result that long term unemployed workers have lower job finding rates (see for example Blanchard and Landier (2002), Hornstein (2012)) so this is not a desirable feature a priori. Thinking of search capital as improving only with success keeps it separate to duration

dependence and its determinants. This doesn't rule out that search capital can be defined in broader terms, allowing for a richer learning process, but narrowing the definition makes it easier to work with. The correlation between number of jobs held and search capital would still hold if workers also learnt from their failed or rejected applications.

The empirical strategy of Section 3.3 relies in this correlation. However how can we know that people that have had more jobs are less productive, or have different preferences to other workers? The interaction of search capital with dual labour markets helps with some of these concerns. For example, in Spain a large share (30%) of workers are employed under *temporary contracts*, with finite duration and low protection in the form of severance payments. The other 70% of the employed hold *permanent contracts*, which have increasing wages and severance with tenure, so these workers little incentive to change jobs after finding permanent employment. Temporary jobs do not immediately translate into stable employment (see for example Güell and Petrongolo (2007); García-Pérez and Muñoz-Bullón (2011)) but instead they often lead to other temporary contracts or unemployment. In other European countries jobs are also unstable for the young, but what sets Spain apart is that even in their late twenties, temporary jobs are common. A possible explanation is that high skill workers enter the market particularly late as well, with the average age of graduation being 27 (OECD (2014)) but engineering, architecture and other technical degrees the average is over 30. This is mostly due to students finishing their degrees after an average of 5 years more than the official time. The result is a growing stock of workers who are frequently searching in the labour market, while simultaneously the security of permanent jobs builds a stock of workers that are very unlikely to ever search again.

Search capital as introduced in this chapter is different from search capital as defined by Carrillo-Tudela and Smith (2017).<sup>1</sup> They refer to the ability of the worker to recall previous employers while employed at another firm, helping them search on the job, while I am referring to the ability of workers to find jobs from unemployment in different firms. This is an important distinction in the empirical model: I do not count recalls back from unemployment as increasing search capital, as the worker doesn't necessarily learn anything by being asked to come back to work at the same firm. Their model is also silent about the implications for unemployment duration, while here it is a central issue.

### **The link to Long Term Unemployment**

As temporary workers are the ones that find themselves more frequently unemployed, during economic expansions search capital is fairly homogeneous across workers. The expansion of 1995-2008 in Spain, together with the job-creation effect of temporary contracts (Güell and Hu (2006)) meant that long-term unemployment fell for all age groups as shown in figure 3.3. This implies that the overall search capital of the unemployed increased during this period. These falling LTU patterns have reversed since 2009, with some authors arguing that the rapid increase in unemployment (and the subsequent increase in LTU) is mainly driven by job destruction from temporary contracts<sup>2</sup>: Firms tend to prefer to reduce their temporary workforce rather than

---

<sup>1</sup>They present search capital as the ability of the worker to recall past employers, so that if current employment ends the worker can go back to their previous employer instead of “falling off the ladder” and start again from a very low productivity job. This has implications for the wage setting process, as firms engage in Bertrand competition to poach workers from other firms. In their model, search capital can also depreciate so that contacts of the worker in their previous job may vanish - so some unlucky workers may not have the option to go back to their past firm.

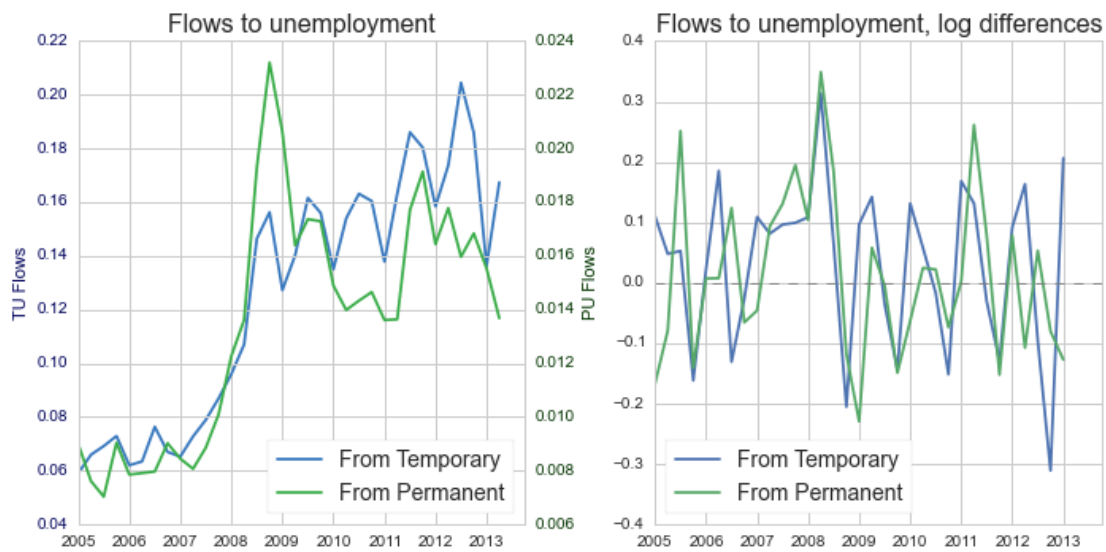
<sup>2</sup>See for example Bentolila et al. (2012) for a comparison between the impact of the recession between France and Spain, where the differences are driven by the wider gap between permanent and temporary contracts in terms of severance payments. Temporary contracts are much easier

Figure 3.3: Long term unemployment in Spain, by age



Source: Own calculations from the Spanish Labour Force Survey (INE (2013))

Figure 3.4: Flows into unemployment, by contract type



Source: Own calculations from INE, *Encuesta de la población activa* (Labour Force Survey), 2013

adjust wages and as a consequence there is too much firing in recessions. The volatility of temporary to unemployment flows does seem to play a significant role in the overall volatility of unemployment (Silva and Vázquez-Grenno (2013)). However the collapse of the construction sector (58% decline in employment from 2008 to 2013), and later a severe financial crisis translated into an increase in lay-offs from both kind of contracts<sup>3</sup>, as figure 3.4 shows. The left panel presents the evolution of flows into unemployment from temporary (blue, left scale) and permanent (green, right scale) contracts. As the difference in scaling shows, separation rates from permanent contracts are about ten times smaller than those of temporary workers. However, the spike of 2008 is higher for permanent contracts (from 0.008 in 2007 to 0.023 in late 2008). The right panel of this figure shows the log difference of these series (log of the flow at time  $t + 1$  - log of the flow at time  $t$ ), where the magnitude of changes in 2008 is again very similar, so in relative terms the increase in job destruction rates from both types of contract was similar. Costain et al. (2010) explained these patterns in their paper as follows: falling productivity thresholds in booms leads to more conversions from temporary to permanent, this creates a growing stock of low productivity permanent workers that are dismissed in a recession.<sup>4</sup> The main driving factor behind the increase in unemployment is not temporary contracts, but high severance payments that prevent firing unproductive permanent workers.

---

to destroy in Spain in their model, so firms react to a fall in productivity by firing most of their temporary workforce.

<sup>3</sup>In fact, following the sovereign debt crisis triggered by these events and the Greece fallout, the government dismissed a significant number of their employees as well. More specifically, real estate lost 22% of its employment, financial services 11%, and 1.7% in the public sector in the 2005-2008 period, according to employment statistics from the Spanish National Statistics Institute (INE (2017)). In 2011 alone, the public sector lost 175,000 workers.

<sup>4</sup>In their model some workers start with high match productivity and thus are promoted to a permanent contract. But stochastic productivity shocks can effectively make them less productive than the hiring threshold. They are kept employed because firing the worker forces firms to pay a lump-sum tax, which for some workers is high enough to keep them in. This is the risk that firms incur when promoting workers, and thus they are more likely going to promote during a period of economic boom.

How does this relate to search capital? Permanent workers have lower levels of search capital because they have been employed longer on average in the same firm, with few incentives to search.<sup>5</sup> When recessions destroy “safe” jobs, there is an influx of bad searchers into the unemployment pool. At the same time, there is more competition for fewer vacancies, so frequent searchers take these jobs first. As these jobs are also predominantly temporary contracts, they flow back into unemployment after not too long, and then again find another job with relative ease. This mechanism depresses the chances of finding jobs for both new entrants with little search experience and dismissed permanent workers. The combination of the market-driven heterogeneity in search abilities of the workforce with a severe recession makes Spain a good choice to test the effects of search capital<sup>6</sup>

However there are other explanations in the literature that can account for the rise in long term unemployment in Spain, although not necessarily for the youth. I highlight the main two competing theories next.

### **3.2.2 Related explanations of LTU**

#### **Unemployment Benefits**

It is a well known theoretical result that a higher unemployment income results in higher unemployment in almost every search model. Consider for example Mortensen

---

<sup>5</sup>For example a worker in the public sector in Spain has to pass an examination process that is very different from private sector application process. In addition to this, public servants have very protected jobs.

<sup>6</sup>It is worth noting though that Spain is not an isolated case and that other Southern European countries with segmented labour markets, such as Italy and Portugal, also saw an increase in their long term unemployment rates. The relatively minor effects in France can be attributed to their milder recession: Search capital increases unemployment when general unemployment is high and there is an influx of bad searchers in the unemployment pool.

(1970): a worker draws a wage offer from a given distribution, then she decides whether to accept the job or to reject the offer and keep searching. The worker sets a reservation wage strategy which depends positively on their unemployment income. She internalizes that she is going to be unemployed for longer in return for a higher future wage. Being richer makes the worker more selective. This is a mechanism that drives more sophisticated models such as Kitao et al. (2017).

The empirical literature seems to confirm these patterns: Lalive (2007), and Krueger and Mueller (2010) find longer periods of unemployment benefits (UB hereafter) results in longer spells of unemployment. Krueger and Mueller (2010) find that time devoted to search increases as the date of benefit exhaustion approaches, but then it is reduced drastically. This implies that although longer UB entitlements can lead to longer unemployment spells they appear to keep the unemployed searching for work. Wadsworth (1991) similarly finds that UB recipients are more attached to the labour market. It seems to be the case in the literature that entitlement (how long benefits last) is more important than the quantity of benefits. This also appears to be consistent with the fact that Northern European countries, where workers are given their benefits in a block payment, have lower unemployment durations overall.

An overly generous benefit can thus lead to longer unemployment durations as an equilibrium outcome. Because the quantity and duration of unemployment benefits are usually dependent on past wages and job durations workers coming from longer durations are expected to take longer to find jobs. The increase in long term unemployment rates could be explained by high-income/wealthy workers choosing to wait for a better job.

In Spain, regular workers tend to have, on average, better paid jobs than temporary workers (because of seniority wage rules and better unionisation) but also because if they are laid off they receive severance payments that, in some cases, can be quite substantial. As more permanent workers have been dismissed in the recession, the composition of the unemployment pool has shifted towards richer individuals.

The generosity of benefits in Spain is high but in line with other European countries. For example, Stovicek et al. (2012) do a detailed comparative study of unemployment insurance and benefits across EU countries and Spain does not stand out in any dimension. While replacement rates are on the high end (approximately 70% on average<sup>7</sup>) the lack of other social benefits, such as housing assistance, child supplements or minimum income, means that Spain is not a particularly generous country for unemployed workers. The maximum extension of UB is two years, similar to France, Portugal or the Netherlands. The median unemployment compensation is 636.22 euros a month compared to the median wage of 1351.72 for temporary workers and 1446.75 for permanent workers.<sup>8</sup> In response to the rapid increase of LTU, Spain did not extend entitlement periods as many US states opted to do. Exhaustion of benefits became commonplace during the recession with close to 50% of all unemployed not receiving any benefits as of 2012.<sup>9</sup> It is hard to argue that the rise in LTU is driven by workers preferring to remain unemployed for a period of time beyond the exhaustion of their benefits.

---

<sup>7</sup>Stovicek et al. (2012)

<sup>8</sup>These are based on calculations using the information in the tax data of the Social Security Working Lives Sample, once the bottom and top 1% percentiles have been removed. See Data section for a detailed account of how these were calculated.

<sup>9</sup>Based on the Spanish LFS and matching Administrative data.



While the generosity of unemployment benefits is a good explanation for the increase in LTU in general, it is not a good explanation for the increase in youth LTU as their lower wages and shorter tenures imply that many of them are not receiving unemployment benefits - and where they are, these are modest and only for a short period.

### **Human Capital Depreciation**

A popular explanation of long term unemployment increases during a recession is that technology shocks can produce redundancies that lead to an immediate and persistent deterioration of productive human capital. This makes it harder for those affected to find subsequent employment. Ljungqvist and Sargent (1998) called this turbulence.

In a more recent paper, Ljungqvist and Sargent (2008) present a model in which, upon losing their job, some workers suffer a sudden and permanent loss of human capital. This leads to lower expected future wages and search effort. Combined with a generous unemployment benefit, individuals who suffer these human capital shocks are discouraged from searching for a new job, leading to long term unemployment. In a similar way, Carrillo-Tudela and Visschers (2013) look at mismatch across occupations and find that most unemployment generated during recessions is what they call “resting” unemployment - workers looking for a job in their previous occupation instead of switching careers. These workers prefer to wait in unemployment in their occupation-specific job market during a recession, in the hopes that their human capital doesn’t fully deteriorate, leading to longer durations of unemployment.

This sudden loss of human capital, it can be argued, is driven by idiosyncratic

shocks to labour demand. For example, in Spain the collapse of the construction sector left many workers unemployed and with a set of skills which is no longer desired by firms. Related industries like building material providers, real estate and financial services also suffer major job losses. More importantly the budget readjustment of 2011 meant a considerable shrinkage of public sector employment.

In this case, the end of a long term job sees part of the human capital of the worker vanish, leading to subsequent job losses. This has been well documented in the displaced worker literature (see Jacobson et al. (1993) and Couch and Placzek (2010) for updated results). If the worker is also entitled to high unemployment benefits then she may be discouraged to search. This mechanism can't fully explain how more experience of temporary contracts (or recent unemployment spells) lead to shorter durations unless temporary contracts increase a worker's human capital more so than a stable contract. However Dolado et al. (2012) have argued that there is no incentive to invest in human capital for temporary workers, documenting a lower incidence of on-the-job training provided by firms compared to permanent workers. In this way permanent contracts could incentivise firm-specific human capital investment while temporary contracts improve transferable skills, leading to observed shorter unemployment spells for those with temporary contracts. Lazear (2009) proposes a model where workers choose to specialise in different kinds of skills depending on how likely an exogenous lay-off can happen, this leads to diversification of human capital in those industries/occupations where jobs are more unstable.

Crucially, the depreciation of human capital can't explain why long term unemployment has risen so dramatically among young workers. Kitao et al. (2017) argue that higher minimum wages are to blame, but then why are some young workers

finding jobs much faster than others? The proliferation of temporary contracts and apprenticeships does not imply that minimum wages are too high, but shows that minimum wages are easily circumvented.

A related issue is the depreciation of human capital *during* unemployment. This could induce *negative duration dependence* - lower exit rates the longer a worker is unemployed. Note that search capital does not decrease with unemployment duration, as workers do not lose any of their search skills while using them to pursue jobs. Therefore this explanation does not compete with search capital.

### **3.3 Empirical Analysis**

Search capital is not observable. Ideally one would like to measure search skills like networks or application strategy. However we may be able to identify skilled searchers by looking at their search outcomes: those who find better jobs faster, once controlling for all other observables, are likely to have better search skills. But these skills could be an inherent trait of the individuals, some people may be born with a natural advantage over others when it comes to finding jobs, having well-connected family or friends for example. This doesn't mean however that people don't get better at searching with time. In order to try to investigate this one would need panel data, with enough variation across time and individuals to see if workers do get progressively better at searching. Working histories datasets, provided by certain social security administrations, are fit for this purpose.

#### **3.3.1 The Data**

The Spanish Social Security administration provides this information from 2004, releasing a sample of close to a million random observations each year. This is the

Muestra continúa de Vidas Laborales (MCVL) analysed in Chapter 1. The data follows individuals through time, adding new observations for the ones dropping out (workers retiring or dying) keeping the sample representative from year to year. Specifically, it consists of a sample of 4% of the working population. The condition to be included in the sample is to have been affiliated with Social Security (either by working, receiving a public pension or being registered as unemployed) in the year of the publication of the dataset. After that year, the MCVL follows the same sample of workers over time, adding new observations each year to replace absences while keeping the sample representative of the population. The MCVL comprises all of the job spells, unemployment spells and retirement periods that are registered by the administration for each individual in the sample. It contains information on personal characteristics (age, gender, date of birth, highest education attained) from the census (last wave dating to 2011), some firm information (size, location, tax code) and information on the job such as industry, occupational scale and type of contract. It keeps track of changes of contract and changes in relation to social security (for example from unemployed to retirement). The MCVL also has information on the self-employed.

The Spanish Social Security also provides a complementary dataset with income tax information, which can be linked to the working histories files via fiscal identifiers. This way it is possible to obtain detailed wage information for most jobs in the sample. It also holds records on severance payments, food coupons, dividends and any other form of transfer between the firm and the worker as payment for work services.<sup>10</sup> Unemployment subsidies received in the last year are also recorded,

---

<sup>10</sup>There is an alternative way of matching jobs with wages through the contribution basis file. Contribution basis are the basic salary threshold that the Social Security uses to calculate contributions to the system. In a recent paper Bonhomme and Hospido (2017) use this data to analyse earnings. I use the tax information because it captures overpay and payments in kind, which

making it possible to approximate the amount of unemployment benefits received in the unemployment spells of the previous year. If the worker has several unemployment episodes in the year for which she received unemployment compensation it is not possible to separate them. However, these occurrences are rare as most unemployed workers can't accumulate enough working spells to be eligible within the year.

I use the 2005-2013 waves of the MCVL. The data contains past information and it would be tempting to use it to go back in time and approximate the workforce of previous years. However, as we go back into the past, the sample stops being representative as it turns younger and only those active in the present are represented in the past.

One concern that arises when using administrative data to study unemployment is that administrations only count registered unemployment spells. For those who can't or choose not to claim UB, the sample only has gaps. However with minor adjustments (using official definitions and labour laws) it is possible to reconstruct the unemployment series to be close to those coming from the Labour Force Survey: figure 1.7 in Chapter 1 showed that without any modifications, the Social Security quarterly unemployment rate is well below the LFS. Adding (1) the days in between finished jobs (2) uncompleted spells as of 2013 and (3) gaps between employment spells for those without the right to claim UB generates the 'MCVL expanded' series<sup>11</sup>, which closely follows the LFS except from the 2005-2008 period. Here the difference is likely coming from the failure of the LFS to adequately capture short unemployment spells for young workers (see Chapter 1). The expansion of the MCVL

---

have increased since the onset of the recession. These kind of payments are more common among temporary and part-time workers.

<sup>11</sup>In practice, there are more restrictions on which gaps between employment spells are added, such as excluding recalls to the same firm. See Chapter 1 for more details.

ensures that spurious spikes in duration at the termination of benefits are ruled out. These can have important effects in the estimation of duration models, as noted in Chapter 2 and Alvarez et al. (2015). There is also an issue around how the LFS classifies unemployment and non-participation, as the flows to employment for both groups are very similar for young workers. This suggests that the unemployment data from the LFS is very sensitive to the definition of unemployment, and thus that in Spain there is little evidence of the non-participants of working age to being completely detached from the labour market. In fact it is well documented that previous changes in this definition caused important breaks in the series. The MCVL is helpful for dealing with these breaks in the LFS data.

Table 3.1: Descriptive Statistics

	count	mean	std	min	25%	50%	75%	max
Weeks	703,061	31.66	42.01	2	6.57	16.29	38.71	462.86
Temporary contracts	703,061	6.29	6.29	0	2	5	8	251
Permanent contracts	703,061	3.79	5.72	0	0	2	5	232
$wage_{t-1}$ (euros, annual)	703,061	19879.72	331783.65	0.01	10334.04	15139.48	19817.51	256319430.50
$wage_{t+1}$ (euros, annual)	602,510	14610.17	63325.83	0.03	4016.85	12578.33	18066.59	28970452.81
$UB_t$ (euros, annual)	694,728	6440.86	11417.19	0	2482.65	5414.58	9086.54	2315531.94
Tenure (years)	703,061	0.97	2.19	0	0.08	0.27	0.82	39.27
Experience (PC, years)	703,061	4.19	5.78	0	0	1.69	6.18	39.48
Experience (TC, years)	703,061	2.42	2.23	0	0.63	1.87	3.59	24.59
UB entitlement (months)	703,061	4.31	7.28	0	0	0	6	24.33
Age	703,061	32.75	8.90	20	25	31	39	54
Male dummy	703,061	0.56	0.50	0	0	1	1	1
Quit dummy	703,061	0.16	0.37	0	0	0	0	1

Source: MCVL, 2005-2013 waves. The sample is all completed unemployment spells, ending in employment, with wage information for the next job, recalls and transitions from self-employment excluded. Wages and unemployment benefits are taken from the fiscal annex of the MCVL (2005-2013).

Table 3.2: Variables of interest

Effect	Control variable	Variable name
Unemployment benefits (quantity)	Observed daily benefits	UB
Unemployment benefits (duration)	Claim dummies for 3,6,12,18,24 months	claim3, claim6, ...
Severance Payments	Indicator for last permanent contract	Last P, tenure
Quality of previous matching	Observed past daily wages	log(past wage)
Quality of future matching	Observed future daily wages	log(wage)
Human Capital (general)	Work experience over potential experience	R. Experience
Human Capital (specific)	Tenure in the last job	Tenure
Search Capital (gain)	Number of temporary contracts held	No. T
Search Capital (atrophy)	Years since last unemployment spell	YEmp

### 3.3.2 Identification Strategy

The goal of the empirical model is to test for a relationship between past and present search outcomes, or how workers with different working histories but similar in all other ways can have different unemployment spells. In particular, do people who have found different jobs in the past have shorter unemployment spells than others with fewer jobs in the past? And if so, are the jobs they find better on average?

My main identification strategy is to use the number of temporary contracts held in the past as a proxy for search capital. There are several reasons for using this approach: first, it can be interpreted as the number of previous successes the worker has had finding a job with different firms. This is important because it is not searching per se that is of interest, but successful search. They indicate that the worker has

had experience of search in the past and that she was good enough to obtain a job. Having worked in different firms also signals adaptability and it can be thought of as an indicator of transferable skills. This is what makes temporary jobs a better indication of search capital than permanent jobs, which usually imply that the worker won't search for a job for a long time. There are some caveats: the firm may "recall" the worker after they have been unemployed for some time, making unemployment an agreed holiday. This is especially true for very frequent jobs, sometimes lasting a day at a time. Using fiscal firm identifiers I exclude recalls and count temporary contract roll-outs (two or more temporary jobs in a row by the same firm) as only one job. This is to ensure there was search involved in the process of finding the next job.

An alternative measure can be the number of past unemployment spells. However, this would rule out search on the job. As I am aiming at capturing how up to date the worker is, these transitions can't be ignored. I also add the number of permanent contracts, but having had a permanent contract in the past can have negative connotations for future employers: to leave a permanent position the worker had to quit (low commitment) or the firm had to lay off the worker (low productivity). Temporary contracts do not have these negative connotations as they just end, and given the low rate of promotion this does not signal much about the worker. Some sectors have seasonal demand upswings (e.g. tourism and agriculture) for which temporary contracts are a natural choice, again having many temporary jobs in these sectors doesn't signal anything negative about the worker. I include both previous temporary and permanent jobs in my regressions, and as expected I find them to have different effects.

A more interesting question is how to capture search skill depreciation. Using



time since last unemployment spell is a good proxy, but as argued before it also means ignoring on-the-job search. Tenure then becomes the best proxy variable, but using this it would be impossible to untangle how much of the effect is due to loss of search skills and how much to loss of specific human capital. Temporary contracts can still capture some of it: consider two workers with the same duration at their previous job, but where one of them had several temporary contracts before that job. The difference in unemployment duration will then be attributed to different search abilities, not to human capital.

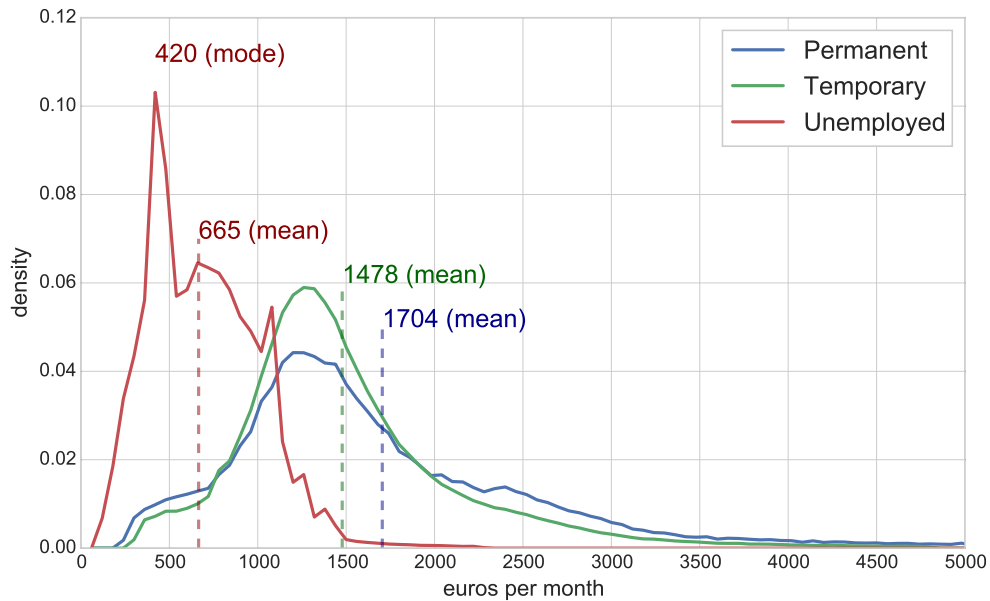
Another variable that is closely related to this is unemployment entitlement period. It is a well know result that there are spikes in job exiting rates close to the expiration of benefits (Card et al. (2007)). This is partly a reflection of employment agencies only recording unemployment while benefits last, as is the case in Spain. These spikes are closely related to the amount of time since the worker was last unemployed, which is another reason to treat the time since last unemployment ( $YEmp$ ) variable cautiously. I include dummies for 3, 6, 12, 18 and 24 months (denoted by the vector  $CLAIM$ ) as these are the most common spikes to control for unemployment entitlement period.

It is also important to have good control variables for the other factors detailed in section 2 that could also be related to past working histories but do not relate to search capital directly: human capital, incentives to search and matching factors. The most important of these controls is taken from the tax information file: past wages, unemployment benefits and severance payments. Past wages are related to productivity, both specific to the individual (more productive workers are likely to have higher wages on average) and to the match itself.

I identify worker-firm pairs using their tax codes, ensuring I correctly identify the previous job. A caveat is that although the tax file only records annual wages, if the worker has more than one job with the same firm in the same year I assume the wages are constant among jobs within the year in the same firm. As I am excluding recalls from the sample, this should not make a significant difference. Some firm-worker pairs can't be identified because some firms have special fiscal identifiers (such as public administrations). The information added by this variable outweighs the potential concerns about noise and missing information, so it is included in all regressions.

I also observe the amount of unemployment benefits the worker receives while unemployed, which is directly related to past wages. As some workers are not entitled to any benefits (quits, self-employed, those who have not accumulated enough tenure) this variable can be interpreted as the effect of more generous benefits on duration, keeping previous wages constant - and keeping duration of benefits constant too. The only caveat is that unemployment benefits are recorded annually and do not distinguish between different unemployment spells. By linking the tax dataset to the working histories I count the days of unemployment within a year in which the worker was eligible to receive them and divide the total amount in euros by the days of unemployment. This measure is likely to be noisy - much more so than wages - so in some regressions I choose not to include it. In the results it can be seen that unemployment benefits don't seem to add much more information than wages and unemployment entitlement dummies. Figure 3.5 shows the resulting benefits and wages distribution.

Figure 3.5: Wage distributions



Source: Own calculations from MCVL, 2005-2013 waves, fiscal annex

Finally, the tax dataset records severance payments separately from general wages as they are excluded from income tax computations - giving an incentive to the worker to report them truthfully. However, the data on severance payments is very noisy (only identified for 2% of permanent workers and 5% temporary workers). Given that most permanent jobs end in a dismissal (as opposed to a quit), a dummy taking the value 1 when the last job was permanent should suffice to control for severance payments, but I add the declared severance payments in the robustness check.

Other variables of interest are the controls for human capital, simply in the form of accumulated job experience in years<sup>12</sup> and past tenure for non-transferable human

---

<sup>12</sup>In alternative specifications I rescale this variable as years of employment over years in the labour force. The results don't change substantially but affect the coefficients for age. I choose to retain age in the regressions, so I revert to use total years of experience.

capital. The usual controls for age, gender and educational levels (split in 4 levels<sup>13</sup>) are included. I also include controls for industry at the one digit-level, province of residence and a proxy for occupational level.<sup>14</sup>

### 3.3.3 Empirical approach

Given the data and the variables, the empirical strategy is to regress the logarithm of weeks of unemployment against the explanatory variables discussed previously plus control variables for individual characteristics, as equation 1 shows. These include age (with a quadratic trend), gender, nationality, industry, region (at the province level) and occupational level. I also include dummies for part-time, collective dismissal and quits. Finally I include two specific dummies for workers coming from construction jobs: one after and one before 2008, the year of the collapse of the building sector.

$$\log(\text{weeks}_t) = \beta_0 + \beta_1 \text{No.T} + \beta_2 \text{YEmp} + \gamma \text{CLAIM} + \beta_3 \text{LastP} + \beta_4 \text{Ten} + \beta_5 \text{Exp} + \beta_6 \log(\text{wage}_{t-1}) + \beta_7 \log(\text{UB}_t) + \delta X + \epsilon \quad (3.1)$$

The unit of observation is a complete unemployment spell: an unemployment spell that ends in a job or self-employment. I exclude self-employed that have lost their job to allow for an easier interpretation of the dummy for last permanent

---

<sup>13</sup>Base category being no formal education/finishing at most primary school, secondary education, pre-university education (including vocational training that requires a secondary education degree) and university education.

<sup>14</sup>Practically all formal jobs have their wages determined by collective agreements at the industry and/or regional levels. These agreements specify different lower bounds for salaries levels within the firm according to professional categories - manager, engineer/skilled analytical worker, high and low skill white collar jobs, etc. The combination of industry-professional category can provide a proxy for occupation, therefore I add these professional categories as controls. García-Pérez and Muñoz-Bullón (2011) argue that they provide a more accurate measure of skill requirement for the job.

contract. In this way, each individual could be appearing multiple times in the regression, but still the sample will be representative of the unemployment pool: some workers find themselves more often unemployed while others rarely appear in the sample. As the aim is to capture the competition of different workers for jobs, the “one spell, one observation” approach is suitable.

A possible concern is that individuals that appear as unemployed multiple times have some unobserved characteristic driving them back and forward from unemployment. Many of these seemingly quick employment/unemployment spells are driven by recalls: a worker returning to the same firm after a brief period of unemployment. I have restricted to sample to consider only workers that change firms to avoid recalls, so this frequent unemployment spells have to come from workers switching jobs. Finally, to address possible unobservable heterogeneity, I redo the exercise with individual fixed effects. Here the interpretation of the coefficients is different: in the pooled sample,  $\beta_1$  would be the marginal effect of having had one more temporary contract in the past on log weeks in unemployment (percentage increase in weeks) across workers. In the fixed effects regressions it would represent the effect of one additional temporary contract on the difference in duration of unemployment spells across time for a single worker. That is, if it is positive (negative) then as the worker accumulates temporary contracts her unemployment spells get longer (shorter) over time. In this way the panel regression measures the effect of accumulating search capital over time, while the pooled regressions measure the effects across workers - some may be born with higher stocks of search capital than others.

The empirical strategy is more simple than in García-Pérez and Muñoz-Bullón (2011), but the variable of interest is different: here the interest lies in identifying

systematic differences in completed unemployment spells, while García-Pérez and Muñoz-Bullón are concerned with estimating hazard rates for exiting unemployment - and ultimately whether or not they have an effect on future upgrades to permanent contracts.

Finally, to complement the analysis above I run a logistic model where I consider the probability that an unemployment spell will last more than a year ( $LTU_1$ ) and two ( $LTU_2$ ) against all the previous variables. As the average spell in Spain is close to a year, I use the two year mark to signal long term unemployment more effectively. The influence of the benefit claiming period (*claim12* and *claim24*) will be clear here. But more importantly, given the increase in long term unemployment in Spain during the recession, skilled searchers having even a small advantage in finding a job could protect them from very long unemployment spells during recessions. These probabilistic regressions consider the impact that one additional success can have on that outcome (long term unemployment).

Of course, it could be that workers with more temporary contracts happen to find jobs faster but these jobs are otherwise worse than the rest. I test for this in two ways: First I regress equation 2 using the information on the next job wages - excluding those who become self-employed.<sup>15</sup> Given that my wage data comes from annual sources and for some jobs it can be imprecise, I restrict the sample to workers who find a job that lasts at least 3 months. Unemployment duration is included as an explanatory variable on its own. This way any remaining effects of all of the other variables are to be interpreted as their independent effect on wages aside from its indirect effect on duration. Consider the case where the coefficient of temporary

---

<sup>15</sup>Not many do in the sample.

contracts is negative and significant in both duration and wage regressions. Then temporary contracts will be correlated with shorter unemployment spells, but also with lower wages. Its overall effect on wage will depend on how unemployment duration impacts wages: its effect on duration may be strong enough to have a positive wage impact overall.

$$\log(wage_{t+1}) = \beta_0 + \beta_1 No.T + \beta_2 YEmp + \gamma CLAIM + \beta_1 No.T + \beta_3 LastP + \beta_4 Ten + \beta_5 R.Exp + \beta_6 \log(weeks_t) + \beta_7 \log(wage_{t-1}) + \beta_8 \log(UB_t) + \delta X + \epsilon \quad (3.2)$$

There is also the concern (as discuss earlier in section 3.2) that even if temporary contracts lead to higher wages, these are compensating for the greater uncertainty of temporary contracts. Here it is good to remember that over 92% of the unemployment spells in the sample end in a temporary contract, so the chances of getting a permanent job out of unemployment are very small. It also makes sense that workers already employed in a temporary job have more leverage when negotiating the terms of future employment than the unemployed, so transitions to permanent jobs are not only promotions within firms but across firms too. Temporary contracts could reduce the average duration of future jobs or impact the chances of getting promoted to a permanent contract. To test this hypothesis, I run another regression with two alternative dependent variables: duration of next job (remember that all spells in the sample end in a job) and time until next unemployment spell.

$$\log(weeks_{t+1}) = \beta_0 + \beta_1 No.T_t + \beta_2 YEmp_t + \gamma CLAIM_t + \beta_3 LastP_t + \beta_4 Ten_{t-1} + \beta_5 Exp_t + \beta_6 \log(weeks_t) + \beta_7 \log(wage_{t-1}) + \beta_8 \log(UB_t) + \delta X + \epsilon \quad (3.3)$$

This last variable takes into account not only the next job, but any subsequent em-

ployment spell. This measure could be right-censored, especially for 2012-2013 observations. In the robustness checks I restrict this regression to early unemployment spells only. Equation 3 also incorporates the length of the current unemployment spell to control for any impact long durations could have - weaker bargaining position, discrimination, etc.

### 3.3.4 Results

Results are shown in tables 3.3 to 3.6. The first table has six columns: the first three correspond to a pooled OLS regression (each unemployment spell counts as one observation) while the others are fixed effects panel regressions. The difference among them is the addition of past wage and present UB variables. This is because not all observations have wage information from their previous job, as noted before. This is the reason why the number of observations of columns 2-3 and 5-6 is smaller.

#### Duration of unemployment

The first thing to note is that the number of temporary contracts held in the past ( $No.T$ ) is significant and negative, even when controlling for individual fixed effects. Each temporary contract reduces the unemployment spell by 3% on average, with slightly higher effects (3.2%) once controlling for past wages (column 2) and unemployment benefits (column 3). Recall that the average number of temporary contracts is 4, so the effect of exposure to temporary contracts for the average worker is 12%. The magnitude of the effect is reduced in the fixed effects regressions. A possible interpretation of this is that this is the effect for each worker throughout their working life, whereas the effect on pooled regressions is the effect between different workers. An alternative explanation is that the effect of temporary contracts



Table 3.3: Regressions on Unemployment Duration

	Pooled OLS			Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
	log(weeks)	log(weeks)	log(weeks)	log(weeks)	log(weeks)	log(weeks)
No. T	-0.030*** (0.0038)	-0.032*** (0.0010)	-0.032*** (0.0010)	-0.005*** (0.0013)	-0.005*** (0.0013)	-0.005*** (0.0013)
YEmp	0.004*** (0.0008)	0.003*** (0.0007)	0.003*** (0.0007)	0.003* (0.0011)	0.001 (0.0013)	0.001 (0.0013)
3 months claim	0.154*** (0.0041)	0.224*** (0.0043)	0.226*** (0.0043)	0.143*** (0.0048)	0.179*** (0.0056)	0.181*** (0.0057)
6 months claim	0.148*** (0.0053)	0.228*** (0.0050)	0.231*** (0.0050)	0.165*** (0.0064)	0.203*** (0.0075)	0.205*** (0.0075)
12 months claim	0.109*** (0.0083)	0.178*** (0.0074)	0.181*** (0.0074)	0.170*** (0.0110)	0.200*** (0.0126)	0.203*** (0.0127)
18 months claim	0.078*** (0.0111)	0.143*** (0.0103)	0.148*** (0.0103)	0.160*** (0.0171)	0.178*** (0.0195)	0.181*** (0.0196)
24 months claim	0.011 (0.0127)	0.046*** (0.0132)	0.049*** (0.0132)	0.113*** (0.0236)	0.110*** (0.0270)	0.113*** (0.0272)
Last P	0.111*** (0.0070)	0.040*** (0.0046)	0.041*** (0.0046)	0.084*** (0.0050)	0.039*** (0.0061)	0.040*** (0.0061)
Tenure	0.011*** (0.0013)	0.016*** (0.0012)	0.016*** (0.0012)	0.020*** (0.0019)	0.027*** (0.0023)	0.027*** (0.0023)
Experience	-0.010*** (0.0005)	-0.007*** (0.0005)	-0.007*** (0.0005)	0.037*** (0.0029)	0.041*** (0.0034)	0.042*** (0.0034)
No. P	-0.038*** (0.0023)	-0.035*** (0.0022)	-0.035*** (0.0022)	-0.008** (0.0029)	-0.008* (0.0032)	-0.008* (0.0032)
age	-0.013*** (0.0023)	-0.002 (0.0017)	-0.001 (0.0017)	-0.038*** (0.0048)	-0.035*** (0.0054)	-0.035*** (0.0054)
log(past wage)		-0.081*** (0.0014)	-0.081*** (0.0014)		-0.043*** (0.0017)	-0.044*** (0.0018)
log(UI)			-0.002*** (0.0001)			-0.001*** (0.0001)
Constant	1.074*** (0.1547)	1.203*** (0.2370)	1.169*** (0.2387)	0.909** (0.3275)	0.993* (0.4014)	0.971* (0.4024)
<i>Controls</i>						
Years	✓	✓	✓	✓	✓	✓
Industry	✓	✓	✓	✓	✓	✓
Occupation	✓	✓	✓	✓	✓	✓
Region	✓	✓	✓	✓	✓	✓
Observations	587,222	465,832	461,369	587,222	465,832	461,369
Adjusted $R^2$	0.546	0.559	0.561	0.462	0.457	0.458
$AIC$	1,502,926	1,189,482	1,176,424	1,082,259	840,389	829,077

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm), self-employed and spells shorter than 15 days.

is reduced in the FE regressions because most of the ability to search is specific to each individual. Gaining more jobs does not greatly improve her chances of finding a job after controlling for individual effects. This would make the differences in search ability across workers more persistent.

The impact of temporary contracts on duration could be non-linear: after a certain amount of contracts the effect could turn positive. Table C.2 in Appendix C shows the results after adding a quadratic term for the number of contracts. It doesn't make a substantial difference for most of the sample: it requires more than 100 temporary contracts for the effect to turn positive, 99% of observations lie outside this range. Another possible caveat is that the effect of temporary contracts may vary for different industries. Table C.3 in Appendix C shows the results of the regression in column 2 (pooled OLS, controlling for past wages) for each industry in the next job. The coefficient on temporary jobs is always significant and negative, suggesting these results hold true in general.

The effect of the length of UB entitlement in months is large. All of the claim dummies are significant and positive. However, the biggest effect is not at the maximum entitlement period (24 months, *claim24*), but at 6 months in the pooled regressions and between 6 and 12 months in the panel regressions. Being entitled to between six months and a year increases time in unemployment by 22.8% (after controlling for unemployment income, column 2). In the fixed effects regression, the effect is smaller, between 16 and 20%. This again may reflect that different people respond differently to the entitlement period of unemployment benefits. Once we remove that unobserved characteristic, the increase is smaller. Compared to the effect of unemployment income ( $\log(UB)$ ) it is clear that it is duration, not size of

the benefits that drives longer unemployment duration.

On the other hand,  $YEmp$  or years since last unemployment spell is significant at the 1% level and positive in all regressions. Controlling for fixed effects also reduces the size of its coefficient. The effect of tenure is greater (an increase of between 1.1 and 1.6% for each year in the last firm). A possible explanation is that a worker that keeps herself employed by different firms, making job-to-job transitions has her search skills relatively up to date. A worker who has “settled” in one company may not have any incentives to search outside the firm (either because of increases in job security, wages or both) and thus her search skills deteriorate faster. This interpretation is appealing since the other channels through which tenure may have an effect on unemployment duration are: benefit entitlement (already controlled for by *claim*), higher past wages (a variable on its own) and severance payment (only for those with a permanent contract, controlled for by *Last P*). The only explanations left are the search channel or a matching effect: since a long past tenure suggests a good match, the worker could be waiting for another good job.

*Last P* itself has a small but negative effect in the pooled regressions when controlling for wages and UB size. However in the fixed effect regressions it has a significant positive effect, extending unemployment duration by between 4 and 11%. This may be interpreted as the contract type being used as a signal of higher skill to employers. However the effect of higher skill could be partly controlled for by adding wages, although workers don’t usually reveal their past wages (at least in their CVs) whereas the type of contract is something that can be easily verified. So across competing workers, having had a permanent contract helps them find a job. When considering the effect across time for individual workers in the FE regressions, the

effect is clearly positive. This likely relates to matching in a similar way to tenure: the worker may not be willing to engage in temporary contracts again, but prefers to wait longer for a permanent position.

The rest of the variables have the expected signs. The only exception is the size of annual unemployment benefit (in logs,  $\log(UB)$ ) which is negative in pooled OLS. A 1% increase of UB *decreases* the length of unemployment by 0.2%. As discussed previously, it is the length, not the size of benefits that matters for durations. The amount of benefits a worker gets are related to their previous wage, so it could be that higher wage earners are also more skilled. However since wages themselves are being controlled for by  $\log(\text{past wage})$ , the remaining effect could be interpreted as people with family responsibilities (who get extra money) leaving unemployment slightly faster. In the fixed effects regression the sign becomes positive - although it is very small. This also signals that once these differences across individuals are removed, more UB means longer unemployment spells, but the effect enters mainly through length rather than the amount paid.

Other variables like the quadratic term on age, the gender and industry dummies can be seen in Appendix C, table C.1. There is a significant positive effect of being previously employed in construction (compared with the base of being employed in the agricultural sector) and the recession makes this effect even higher.

To complement the previous analysis, table 3.4 shows the result of the logistic regression on the probability of becoming long term unemployed. The first two columns take into account all unemployment spells, even those unfinished by the end of 2013. This could lead to misleading results for unemployment spells starting in

Table 3.4: Probability of Long-Term Unemployment

	All Spells		Finished Spells	
	(1) $P(\geq 1year)$	(2) $P(\geq 2years)$	(3) $P(\geq 1year)$	(4) $P(\geq 2years)$
No. T	-0.087*** (0.0011)	-0.125*** (0.0020)	-0.090*** (0.0015)	-0.135*** (0.0031)
YEmp	0.007*** (0.0013)	0.018*** (0.0016)	0.003 (0.0018)	0.015*** (0.0027)
3 months claim	0.481*** (0.0110)	0.478*** (0.0163)	0.481*** (0.0144)	0.574*** (0.0239)
6 months claim	0.780*** (0.0115)	0.656*** (0.0167)	0.830*** (0.0149)	0.770*** (0.0246)
12 months claim	1.045*** (0.0153)	0.878*** (0.0209)	1.063*** (0.0202)	0.985*** (0.0314)
18 months claim	1.051*** (0.0196)	0.789*** (0.0260)	1.047*** (0.0272)	0.935*** (0.0412)
24 months claim	0.498*** (0.0256)	0.154*** (0.0352)	0.642*** (0.0364)	0.605*** (0.0569)
Last P	0.238*** (0.0104)	0.227*** (0.0148)	0.213*** (0.0134)	0.179*** (0.0213)
Tenure	0.038*** (0.0022)	0.016*** (0.0028)	0.041*** (0.0031)	0.031*** (0.0045)
Experience	-0.015*** (0.0008)	-0.015*** (0.0011)	-0.014*** (0.0011)	-0.024*** (0.0018)
No. P	-0.146*** (0.0038)	-0.195*** (0.0060)	-0.124*** (0.0053)	-0.174*** (0.0097)
log(past wage)	-0.211*** (0.0035)	-0.229*** (0.0051)	-0.266*** (0.0043)	-0.333*** (0.0065)
log(UI)	-0.002*** (0.0003)	-0.003*** (0.0004)	-0.002*** (0.0003)	-0.003*** (0.0005)
age	-0.018*** (0.0031)	-0.044*** (0.0043)	0.022*** (0.0041)	0.032*** (0.0066)
Constant	-0.379*** (0.0773)	-0.766*** (0.1137)	0.790 (0.5222)	0.614 (0.6264)
<b><i>Controls</i></b>				
Years	-	-	-	-
Industry	✓	✓	✓	✓
Occupation	✓	✓	✓	✓
Region	✓	✓	✓	✓
Observations	726,839	726,839	461,369	461,369
<i>AIC</i>	603,127	327,490	369,427	173,783

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm), self-employed and spells shorter than 15 days.

2012 onwards because of right-censoring of unemployment spells, so I consider only finished spells only in columns 3-4. The number of temporary contracts is negatively correlated with the probability of LTU, both when we consider LTU as 1 year or more or 2 years or more. In particular, one extra temporary contract diminishes the probability of LTU (as one or more years of unemployment) by 8.3% on average ( $1 - e^{-0.063}$ ), and 12.8% for two years or more. These coefficients increase to 8.6% and 12.6% respectively when only considering finished spells.

The coefficients on the right to claim dummies are all positively, as expected. Having the right to claim for 18 months increases by 2.86 ( $e^{-1.105}$ ) the odds of ending in long term unemployment (over one year), with a similar magnitude for only unfinished spells. However, the quantity of unemployment benefits as measured by  $\log(UB)$  decreases the odds of LTU for finished and unfinished spells. This again points out at the strong effect of unemployment benefit duration compared to the amount workers receive. It is worth noticing that last wages are negatively correlated with probability of LTU for all durations (an increase in past wages by 1% decreases the chances of LTU by 20% for the average worker).

Years since last employment spell are also significant and positive. Its effects are relatively modest (an increase of 0.7-1.8% per extra year since last time employed). Given the small coefficients of this variable in the previous regression, these results suggests that having spent many years away from unemployment has a negative impact in future unemployment, spells, making them longer, as expected. The coefficients of tenure are stronger, suggesting that tenure could be a more robust indicator of search capital depreciation. The main difference between both is that years since last employment spell ignores job-to-job transitions, while tenure resets

with every new job. Finally, the coefficients for work experience are negative and significant, as in the previous table.

Lastly having lost a permanent contract is positively correlated with the chances of ending in long term unemployment. This corroborates the findings on the previous regression. The effects of age are once again quadratic, so although the linear coefficient is negative the square term is positive, indicating that both young people and mature workers are more likely to be found in long term unemployment than middle aged workers.

### **Wages in the next job**

Table 3.5 shows the outcomes of wage regressions, where  $\log(\text{next wage})$  is the natural logarithm of annual wages in the job after the current unemployment spell. Column 1 show the results for all observations, while column 2 only includes jobs that last at least three months and columns 3 only considers jobs that last at least six months. These restrictions ensure that there is less noise in the dependent variable (log annual wages in the next job), at the expense of reducing the sample size.

The first variable is weeks of unemployment spell, which is significant and negative throughout all regressions. This suggests that the longer the unemployment spell, the lower future wages: each 1% increase in weeks of unemployment reduces future wages by between 2.9 and 10.61%. As both variables are in logs, this can be interpreted as a wage-unemployment elasticity. The negative coefficient implies that spending more time in unemployment reduces future wages of workers, so it becomes harder to justify that workers with longer unemployment spells are waiting for a better offer. This implies that variables that were negatively correlated to du-

Table 3.5: Regressions on Future Wages

<i>Sample</i>	(1) log(next wage) all jobs	(2) log(next wage) jobs > 3 months	(3) log(next wage) jobs > 6 months
log(weeks)	-0.1061*** (0.0016)	-0.0350*** (0.0013)	-0.0295*** (0.0015)
No. T	0.0089*** (0.0003)	0.0026*** (0.0004)	0.0018*** (0.0005)
YEmp	0.0023*** (0.0005)	0.0026*** (0.0004)	0.0023*** (0.0005)
3 months claim	0.0669*** (0.0044)	-0.0069* (0.0035)	-0.0150*** (0.0038)
6 months claim	0.1203*** (0.0047)	0.0265*** (0.0038)	0.0077 (0.0041)
12 months claim	0.1479*** (0.0060)	0.0454*** (0.0050)	0.0192*** (0.0056)
18 months claim	0.1447*** (0.0078)	0.0525*** (0.0067)	0.0345*** (0.0075)
24 months claim	0.0726*** (0.0096)	0.0291*** (0.0083)	0.0201* (0.0095)
Last P	0.0191*** (0.0043)	0.0152*** (0.0033)	0.0113** (0.0036)
Tenure	0.0004 (0.0008)	-0.0007 (0.0007)	-0.0011 (0.0008)
No. P	0.0025** (0.0010)	0.0093*** (0.0011)	0.0127*** (0.0013)
Experience	0.0107*** (0.0003)	0.0055*** (0.0003)	0.0039*** (0.0003)
log(past wage)	0.1261*** (0.0018)	0.0575*** (0.0014)	0.0502*** (0.0015)
age	0.0434*** (0.0014)	0.0210*** (0.0011)	0.0178*** (0.0012)
Constant	7.1325*** (0.2263)	8.0257*** (0.2689)	8.0363*** (0.3124)
<b><i>Controls</i></b>			
Years	✓	✓	✓
Industry	✓	✓	✓
Occupation	✓	✓	✓
Region	✓	✓	✓
Observations	465,832	199,679	123,238
Adjusted $R^2$	0.146	0.203	0.250
$AIC$	1,284,934	278,985	135,067

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm) and self-employment



ration have a positive impact on wages by shortening the unemployment spell. The effects of these variables should be interpreted as their marginal effects independent of duration.

One of the variables that both reduces duration and increases wages is the number of temporary contracts. Its coefficient is small (between 0.89% and 0.18%). For a median annual wage of 20,672 euros, an extra temporary contract adds between 183 and 37 euros annually. The small effects suggest that the number of temporary contracts is only weakly positively related to wages, both directly and through its negative impact on unemployment duration. This is an important result because it shows that workers with more temporary contracts (that are better searchers) do not find worse jobs than other workers on average. There are in fact slightly better paid.

Years since last employment ( $YEmp$ ) is significant and positive, suggesting that having spent long times since the last unemployment spell doesn't hurt the job prospects of workers. A possible interpretation could be that workers that have lost their search skills after a long period of employment eventually find similar jobs to other workers, reinforcing the idea that search skills are not the same as productive skills.

Log past wages are highly correlated with present wages, as expected. The coefficients of claiming period dummies become smaller for longer jobs, with 3 month claim coefficient turning negative for jobs longer than 3 months. These dummies represent the effects are over someone without the right to claim UB, suggesting that longer entitlement to unemployment benefits may improve wages via better match-

ing quality or better outside option. The absence of the quantity of unemployment benefits is due to its effects on sample size, as it requires to match two firms and her unemployment benefits with the tax records.

The coefficient for the dummy for permanent contract is positive for all next job durations. The coefficient for tenure is insignificant, suggesting that maybe the effects of longer tenure (higher probability of depreciated search skills) are picked up by *LastP* and *YEmp*. Age and work experience are significant and positive, as expected. I interpret this as evidence of human capital accumulation - aside from individual productivity which is meant to be captures by past wages.

### **Duration of next job spell**

Finally table 3.6 shows the results of the regressions on next job duration. The first three columns are for pooled data (one spell, one observation) and the last three are for fixed effect regressions, as table 3.3 did. Columns 1 and 4 show the regressions where duration of next job in log(weeks) is the dependant variable, while 2 and 4 have the next employment spell as the dependant variable, that is, considering not only how long the next job is but all the subsequent employment spells until the next time the worker is unemployed. This could be a better measure for temporary workers who may gain permanent contracts in subsequent jobs, or may get other temporary contracts. Finally columns 3 and 6 show the results of a logistic regression where the dependent variable is the probability of obtaining a permanent job after the end of the current unemployment spell - which is one if the next job is permanent and zero if it is temporary. Only 8% of all unemployment spells in the sample end in a permanent contract, so this is a rare occurrence. These regressions aim to capture how stable the jobs that are found by workers are.

Table 3.6: Regressions on duration of next job

	Pooled data			Fixed Effects		
	(1) Duration of next job (log weeks)	(2) Duration of next employment spell	(3) $\Pr(P_{t+1} U_t)$	(4) Duration of next job (log weeks)	(5) Duration of next employment spell	(6) $\Pr(P_{t+1} U_t)$
No. T	-0.051*** (0.0007)	-0.013*** (0.0003)	-0.068*** (0.0020)	0.039*** (0.0054)	0.019*** (0.0035)	0.243*** (0.0081)
log(weeks)	0.180*** (0.0029)	-0.006** (0.0020)	-0.016** (0.0057)	0.077*** (0.0040)	-0.000 (0.0021)	0.003 (0.0112)
YEmp	0.016*** (0.0012)	0.008*** (0.0021)	0.018*** (0.0020)	-0.017*** (0.0021)	-0.029*** (0.0024)	-0.007 (0.0058)
3 months claim	0.114*** (0.0079)	0.107*** (0.0063)	0.022 (0.0166)	-0.051*** (0.0105)	-0.006 (0.0064)	-0.044 (0.0292)
6 months claim	0.163*** (0.0091)	0.176*** (0.0087)	0.089*** (0.0180)	0.031* (0.0136)	-0.054*** (0.0094)	-0.003 (0.0366)
12 months claim	0.138*** (0.0134)	0.117*** (0.0153)	0.102*** (0.0248)	0.097*** (0.0220)	-0.107*** (0.0184)	0.148* (0.0580)
18 months claim	0.028 (0.0184)	0.028 (0.0248)	0.038 (0.0324)	0.172*** (0.0331)	-0.045 (0.0327)	0.106 (0.0892)
24 months claim	-0.054* (0.0232)	-0.100** (0.0344)	0.005 (0.0397)	0.333*** (0.0440)	0.181*** (0.0517)	-0.041 (0.1202)
Last P	0.105*** (0.0079)	0.102*** (0.0073)	0.393*** (0.0153)	-0.079*** (0.0112)	0.043*** (0.0081)	-0.626*** (0.0248)
Tenure	0.003 (0.0021)	0.045*** (0.0039)	-0.004 (0.0034)	-0.062*** (0.0040)	0.009* (0.0042)	0.028** (0.0103)
Experience (T)	0.037*** (0.0014)	-0.009*** (0.0010)	-0.053*** (0.0033)	-0.413*** (0.0093)	-0.512*** (0.0079)	0.409*** (0.0257)
Experience (P)	0.016*** (0.0007)	-0.002** (0.0006)	-0.002* (0.0012)	-0.211*** (0.0095)	-0.780*** (0.0102)	-0.177*** (0.0183)
No. P	-0.014*** (0.0021)	-0.012*** (0.0009)	0.293*** (0.0058)	0.019*** (0.0051)	0.084*** (0.0119)	-0.592*** (0.0158)
log(past wage)	0.075*** (0.0025)	0.045*** (0.0018)	0.093*** (0.0056)	-0.032*** (0.0034)	0.015*** (0.0019)	0.024** (0.0094)
log(UI)	0.001*** (0.0002)	0.001*** (0.0001)	0.001* (0.0004)	-0.003*** (0.0002)	-0.000 (0.0001)	0.001* (0.0007)
age	0.035*** (0.0024)	0.020*** (0.0019)	-0.004 (0.0048)	0.266*** (0.0127)	0.236*** (0.0091)	0.032 (0.0268)
Constant	-0.091 (0.3614)	-0.331 (0.3031)	-3.395*** (0.0937)	-2.746*** (0.5832)	-1.853*** (0.3216)	
<b>Controls</b>						
Years	✓	✓	✓	✓	✓	✓
Industry	✓	✓	✓	✓	✓	✓
Occupation	✓	✓	-	✓	✓	-
Region	✓	✓	-	✓	✓	-
Observations	461,300	357,914	472,006	461,300	357,914	77,992
Adjusted $R^2$	0.158	0.133	-	0.063	0.258	-
AIC	1,747,415	1,068,172	284,635	1,373,508	576,323	45,637

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm), self-employed and spells shorter than 15 days.

Looking at the pooled regressions, number of temporary contracts appears to be bad for job stability on average: new jobs are shorter the more temporary contracts you have. This is true when future job duration is restricted to 6 months at least (for consistent wages and to keep very short/menial jobs out of the sample). But when we look at its impact through time in the fixed effects regressions, as the worker accumulates more jobs the duration of next jobs becomes longer too. This suggests that learning to search helps workers get jobs with longer durations as well as higher paying jobs, within individuals. These results are the same for the probability of getting a permanent job out of unemployment. The fact that we need to control for individual fixed effects to see a positive impact suggests that there is an unobserved component that makes some workers more prone to stability. All regressions have controls for occupational level and industry, so this unobservable factor seems to be independent from sectoral composition. Another interpretation is that workers with many temporary contracts are very good at finding jobs, and are less concerned with employment stability - this is not to say that they wouldn't prefer more stable jobs, but that they are willing to accept short jobs more often than workers who are not used to temporary contracts.

This observation, together with the fact that people with more temporary jobs finds jobs faster, suggests a trade-off of waiting for a more stable job versus staying in unemployment for longer. What may be a good strategy when the job market is booming could turn into a higher chance of long term unemployment during recessions: if the jobs available in recessions are worse (as the rise in part-time and shorter TC suggest) being willing to accept an unstable job can keep a worker out of long term unemployment.

Conversely, years since last unemployment spell ( $YEmp$ ) has a positive impact in the cross section, but the fixed effects regressions shows a negative correlation with future employment duration.

Longer unemployment spells ( $\log(\text{weeks})$ ) have an ambiguous effect of future job stability: they are positively correlated to duration of the next job (columns 1 and 4) but mildly negatively with the duration of the next job spell (taking into account future jobs as well). It also reduced the odds of being hired with a permanent contract out of unemployment in the pooled regression (column 3), but it becomes insignificant in the panel regression (6). Past wages and unemployment benefits generally increase next job duration and the odds of being hired with a permanent contract, except in the fixed effects regression on duration of the next job. Age is positively related to job duration, and more strongly so if we consider continuous employment instead of just one job (columns 2 and 5 versus 1 and 4).

### 3.4 Theoretical Model

I have presented evidence of the positive effect of having more jobs can have in future unemployment outcomes, including controls for other potential explanations (human capital, incentives to search and ladder-claiming effects).

Here I present a search model with savings that introduces search capital as a mechanism that affects workers job finding probabilities that evolves through time as the worker learns to search. The goal of the model is to show that the introduction of search capital into an otherwise conventional model helps to explain the evolution

of unemployment patterns through a workers life (particularly in the early years). A standard model with a constant, single job finding rate produces flat flows throughout a workers life. With search capital, the composition of the unemployment pool changes over time, as searchers get better with age.

The model joins two separate strands of the search literature. First, dynamic models with savings and human capital depreciation as in Ljungqvist and Sargent (2008) and Kitao et al. (2017) give the means and motive for older workers to remain unemployed for long. The means are that older workers have more resources to wait for better employment, whereas the motive is the desire to smooth out income shocks derive from the loss of employment (and human capital). In this models the loss of human capital upon job loss means that the jobs they are going to find once unemployed are worse than the job they lost. In this way it is the ‘ladder’ component of human capital that motivates older workers to wait. The second strand is the dual-market literature as in Blanchard and Landier (2002), Güell (2003), Costain et al. (2010) and Bentolila et al. (2012) among others. In these models temporary contracts are modelled explicitly. However, most of these models are focusing on the relation between heterogeneous firing costs and unemployment, so they assume hand-to-mouth, risk neutral workers, which leaves out the consumption smoothing motive. These mechanisms should play a key role in shaping worker’s preferences over temporary and permanent contracts. In their absence, most of the literature imposes the conversion of temporary contracts into permanent ones, or assumes higher wages under permanent contracts.<sup>16</sup>

---

<sup>16</sup>A few notable exceptions to this include Alonso-Borrego et al. (2005) and Cozzi and Fella (2016). The latter shows the effect that risk aversion and consumption smoothing can have in the presence of tenure-increasing severance payments.

In order to model the effects of search capital as workers age, I use a dual-market model that allows workers to save and be risk averse, in the spirit of Cozzi and Fella (2016). In this way, young workers would not have enough savings to smooth out consumption while unemployed so they will accept worse jobs. As they age, if they manage to save more, they will become more selective on their jobs they choose. Here a permanent contract offers not only higher *and* stable earnings. Search capital can alter these patterns, by making experienced young workers more efficient at search and older workers that lose a job after tenure less efficient at search. Which of the two effects dominates drives the results from the model.

Adding dual markets, savings, risk aversion and search capital makes for a rich but complicated model. In order to keep the model simple while retaining elements outlined above, I make some simplifying assumptions on other aspects of the economy. The main assumption is the absence of the firm's problem from the model. Workers draw an offer from an exogenous distribution and accept or reject the job. The assumption of a fixed wage instead of wage bargaining or another wage-setting mechanism may be strong, as it implies little correlation between present and future wages. An alternative would be to introduce some form of general human capital (as in Kitao et al. (2017)). But given how stable permanent jobs are, and because asset accumulation allows workers to increase their reservation wages, the random search assumption is not as strong as it initially appears. In fact it is not too far from match-specific productivity in search and matching models. Moving away from partial equilibrium, the ideal alternative would be to introduce wage bargaining. This would not change the results much, it would only introduce a connection between the external option of workers and unemployment - through less vacancy posting. There is also the concern that the wage distributions from the model would not correspond

to those in the data, particularly for lower wages. Here the introduction of expiring unemployment benefits implies that poor young workers are willing to accept the lowest wage of the distribution.

Another important assumption is the introduction of a no-borrowing constraint that binds for the poorest individuals. In light of the lack of unemployment benefits for a large share of workers as early as 2010. I consider this addition an important feature, and it helps match the observed wage distributions, especially at the lower end. It also gives risk averse workers incentives to save, in order to self-insure against long unemployment spells that may result in them being close to the constraint. The financial aspect of the model is not of primary interest, but a similar framework could easily be adapted to think about financial problems, such as tightening borrowing constraints or housing. The fact that financial constraints are secondary in the model is helped by its being a partial equilibrium model, so households take the interest rate as given. Finally, I will not look into inequality but search capital speaks to it from another point of view: the workers that achieve stability get the best outcomes, but are also exposed to greater risk if they lose their job.

There are other aspects of the model that are not very common: unemployment benefits expire and not all worker are covered by them. This is important to accurately reflect the problems young workers in Spain face, both before and during the crisis: as they have not accumulated enough assets, the threat of low consumption makes them lower their standards for employment such that very temporary, low wage jobs are accepted. As workers build up a stock of assets, they are able to raise their reservation wages and access better jobs.



### 3.4.1 Value functions

Time is discrete, and one time unit corresponds to a month. Workers are risk averse and live indefinitely. Agents can save but can't borrow. They accumulate capital by saving their income in each period, and can be employed with a permanent contract ( $P$ ), employed with a temporary contract ( $T$ ), unemployed with unemployment benefits ( $U$ ) and unemployed without unemployment benefits ( $0$ ). They are born with zero assets and an initial level of search capital  $s_0$ .

Search capital is discrete, increasing with a stochastic probability each time the worker finds a job from unemployment, and decreasing with a stochastic probability each period the worker is employed in a long-term (permanent) job.

#### Employed workers

Permanent and temporary jobs differ in four fundamental aspects: workers in permanent jobs accumulate job-specific human capital ( $h$ ), which increases their wages. Both kinds of jobs suffer a stochastic, exogenous job destruction rate, but it hits temporary jobs more often ( $\delta_T \gg \delta_P$ ). If a temporary job is destroyed, with probability  $\delta_{T0}$  the worker is not eligible for unemployment benefits; permanent workers are always entitled to unemployment benefits after losing their jobs. Finally, permanent workers receive external (temporary) offers with probability  $\alpha_{PT}$  and can choose to leave or stay, while temporary workers get a promotion to a permanent job with probability  $\alpha_{TP}$ . I allow for them to refuse the offer and go back to the unemployment pool without benefits. By design, an entry-level ( $h = 0$ ) permanent job is always preferable to a temporary job with the same wage: in a permanent position wages can only increase and unemployment risk is lower (and always insured).

Permanent workers suffer a stochastic risk of search capital depreciation  $\pi_{s'|s}^-$  each period, while temporary workers are able to keep their search capital. This assumption reflects both that temporary workers change jobs often (so they are able to keep their knowledge “fresh”) and that many permanent workers are likely to retire in the job they get.

This is the continuation value if the worker chooses to stay in her current employment, but I allow workers to quit to unemployment if they find it more profitable. This could be the case if the worker accepts a low wage job when assets are low, but as she builds up capital she decides to search again. I assume temporary quitters are not entitled to benefits, but permanent workers are. The value functions are:

$$V^T(w, a, s) = u(c(w, a)) + \beta \max\{V^0(a', s), \tilde{V}^T(w, a', s)\} \quad (3.4)$$

$$\tilde{V}^T(w, a', s) = \alpha_{TP} \max\{V^0(a', s), V^P(w, 0, a', s)\} + \delta_T V_0^U(a', s) + (1 - \delta_T - \alpha_{TP})V^T(w, a', s)$$

*st.*

$$c + a' \geq (1 + r)a + w$$

$$V^P(w, h, a, s) = u(c(w, a)) + \beta \max\{V^U(a', s), \tilde{V}^P(w, h', a', s')\}$$

$$\begin{aligned} \tilde{V}^P(w, h', a', s') &= \alpha_{PT} \int \max\{V^P(w, h, a, s), V^T(w', a, s)\} dF(w') + \delta_P V^U(a', s) + \\ &(1 - \delta_P - \alpha_{PT}) \left[ p(h)V^P(w, h', a', s) + (1 - p(h))V^P(w, h, a', s') \right] \end{aligned} \quad (3.5)$$

*st.*

$$c + a' = (1 + r)a + w(h)$$

$$s' = \pi_{s'|s}^- s' + (1 - \pi_{s'|s}^-)s$$

Where  $\tilde{V}^j(w, h, a, s)$  denotes continuation value of current employment and apostrophes denote next period variables. No time subscripts are necessary as the four state variables suffice to describe the workers' problem. Search capital depreciation follows a Markov process with downgrading probability of  $\pi_{s'|s}$  and upgrading probability of 0 - the assumption is that any external job offers arrive exogenously and don't contribute to search capital.

Human capital is limited to the firm, and this could be a strong assumption. However it is equivalent to models with full depreciation of human capital upon job loss. I also choose to ignore retirement decisions by assuming a stochastic drop-out rate ( $\beta = \tilde{\beta} + \rho$ ). The reasons for this are that I am more interested in young and middle aged worker's dynamics, as other authors have written extensively over how long-term unemployment interacts with retirement decisions.<sup>17</sup> A straightforward extension would be to set the model in finite time and introduce age-specific shocks, like retirement. This would help match the older workers' unemployment. Asset accumulation still affects the patterns of unemployment over the lifetime via reservation wages, whereby older workers are able to afford to wait longer.

## Unemployed workers

If receiving unemployment benefits agents receive  $b$  and if they run out of benefits they receive zero<sup>18</sup> and have to use their assets for consumption - as a retired person would do. All unemployed agents search for a job each period, and successfully get a job offer of type  $j \in \{P, T\}$ . The arrival rate  $\alpha_j(s)$  is increasing in search capital  $s$ . The job consists of a take-it-or-leave-it entry wage offer  $w$ . If she accepts it,

---

<sup>17</sup>See Kitao et al. (2017), Ljungqvist and Sargent (2008)

<sup>18</sup>Technically they receive 1 unit to prevent them from starving.

with probability  $\pi_{s'|s}^+$  her search capital increases and with complementary probability  $(1-\pi_{s'|s}^+)$  she stays at her current level. The assumption is that only successful searchers contribute to increase workers' search capital. Also if workers learn for failures this would create positive duration dependence, which is not present in the data.

If receiving unemployment benefits, the worker can lose her benefits with probability  $\delta_0$  in each period. The stochastic benefit expiration rate helps to keep the model simple by not keeping track of previous employment history. The stock of search capital is the only history-dependent state variable, besides from assets and human capital for permanent workers.

$$\begin{aligned}
V^U(a, s) = & u(c(b, a)) + \beta \left( \alpha_T(s) \int_0^{\bar{w}} \max \{V^U(a', s), V^T(w, a', s')\} dF(w) + \right. \\
& \alpha_P \int_0^{\bar{w}} \max \{V^U(a', s), V^P(w, a', s')\} dF(w) + \\
& \left. (1 - \alpha_T - \alpha_P)[(1 - \delta_0)V^U(a', s) + \delta_0 V^0(a', s)] \right) \quad (3.6)
\end{aligned}$$

*st.*

$$c + a' = (1 + r)a + b$$

$$s' = \pi_{s'|s}^+ s' + (1 - \pi_{s'|s}^+) s$$

$$\begin{aligned}
V^0(a, s) = & u(c(0, a)) + \beta \left( \alpha_T(s) \int_0^{\bar{w}} \max \{V^0(a', s), V^T(w, a', s')\} dF(w) + \right. \\
& \alpha_P(s) \int_0^{\bar{w}} \max \{V^0(a', s), V^P(w, a', s')\} dF(w) + (1 - \alpha_T(s) - \alpha_P(s))V^0(a') \left. \right) \quad (3.7)
\end{aligned}$$

*st.*

$$c + a' = (1 + r)a$$

$$s' = \pi_{s'|s}^+ s' + (1 - \pi_{s'|s}^+) s$$

## Solving the Model

A solution to the model is a set of reservation wage rules and policy functions for savings such that: (1) workers maximise their utility given their initial states (2) the reservation wages are consistent with the implied value functions.

I solve this problem by value function iteration. The use of first order conditions would be faster, but it is complicated by the “kinks” that result from the discrete choices (accept/reject a job for the unemployed and quit for the employed). The addition of search capital as discrete state variable adds to the dimensionality of the problem. Moreover, as the problem becomes highly non-linear close to the borrowing constraint (low assets/low wages combinations) functional approximation is complicated. All these factors make VFI a slow but safe choice.

Once the policy functions have been found, I simulate the economy from 40 years to find a steady state for the economy and see the evolution of unemployment through the working lives of workers. This also provides moments to compare to the data: average stock of employment/unemployment and flows between states at each age and in the aggregate.

### 3.4.2 Calibration

#### Preferences

The utility function is CRRA, with a risk aversion parameter of 2. Interest rates are set to 2% annual. The discount factor  $\beta$  is set to 0.98.

Table 3.7: Calibration

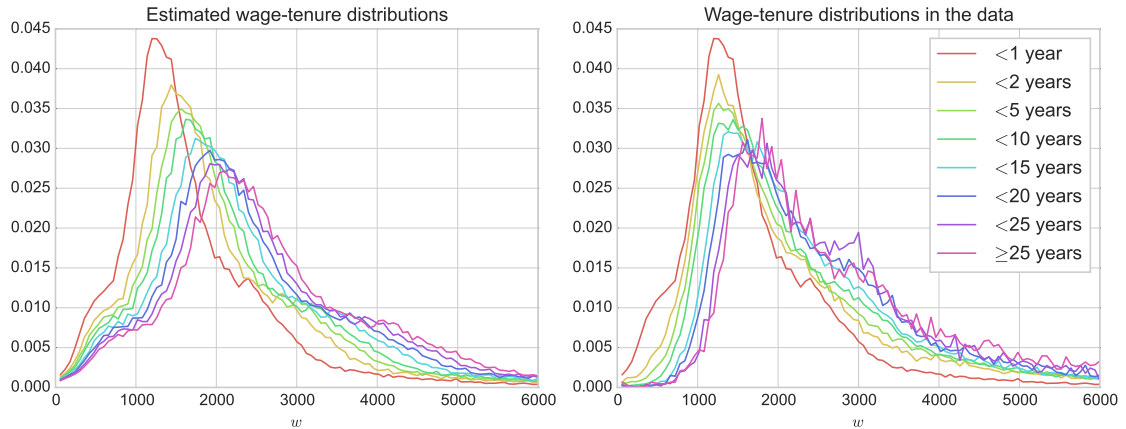
Parameter	Value	Target
$F_T(w)$	-	wage distribution for TCs
$F_P(w)$	-	wage distribution for PCs new hires
$b$	695.52	average UB
$\alpha_T(1)$	0.116	UT transition rates at age 20
$\alpha_P(1)$	0.019	UP transition rates at age 20
$\alpha_{TP}$	0.023	average TP flow
$\alpha_{PT}$	0.226	average PT rates, modified*
$\delta_P$	0.007	average PU flow
$\delta_T$	0.030	average TU flow
$\delta_{T0}$	0.15	average T0 flow
$\delta_0$	0.079	average U0 flow
$p(h)$	-	tenure wage distribution
$s_0, s_1, s_2$	{0.666, 1, 1.666}	duration of unemployment for different NoTs
$\pi_{s' s}^+$	{1, 1, 0}	duration of unemployment for different NoTs
$\pi_{s' s}^-$	{0, 0.0507, 0.0507}	5 year average
$r$	0.0016	2% annual <sup>19</sup>
$\beta$	0.98	
$\sigma$	2.0	Literature

\* consistent with the share of workers at age 20 that accept an average temporary the job offer.

## Wage Distributions

The wage distributions that workers face are taken from the actual wage distributions in the years 2005-2008 as shown in the histogram in figure 3.5. They are binned from 60 to 6000 euros a month and normalised. I let the distribution of wages out of unemployment vary, but when temporary workers get a promotion they start from their temporary wage. To get the human capital coefficients that drive wage increases for permanent workers I look at the evolution of wage distributions on stayers. Here I assume for simplicity a linear wage increase with tenure (so  $w(h) = w * h$ ) and

Figure 3.6: Tenure-Wage distributions



Source: Own calculations and MCVL, 2005-2013 waves, fiscal annex

minimize the distance between the observed distributions. The results are shown in picture 3.6. These assumptions ensure that wage evolution is consistent with the data.

### Employment shocks and job arrival rates

Figure 3.7 shows the average monthly transition rates by age between employment and unemployment, for temporary and permanent contracts. The first thing the plot shows is that there are reasons to prefer a constant job separation rate, at least from temporary contracts (bottom left panel). Job expiration being constant is more difficult to justify but keeping track of past tenure can make the model too large. Simply controlling for age could solve this issue. At present the model relies on the change of reservation wages/job finding rates through time to match young workers unemployment rates.

Getting an estimate of job offers arrival rates is not trivial, as they are a combination of reservation wages and actual job arrival rates. I choose to target job

finding rates at age 20 - the age I take as a “start of the working life” for workers. By assuming new entrants have no assets and no search capital, I can take the job finding rates in the data as the true arrival rate of offers, as these workers accept any wage offer they receive. This ensures reservation wages do not cloud the calibration. In the results it can be seen that this assumption makes the patterns of job finding rates consistent with the data.

Finally the permanent to temporary arrival rate is subject to reservation wage constraints, as not all workers accept a job, not even the very young ones. I take a similar approach by targeting the job switching rate at age 20, solve the model, and then calculate how many permanent workers would switch if offered the average temporary wage. Given this estimate, I update<sup>20</sup> the job offer arrival rate and solve again.

### Search capital parameters

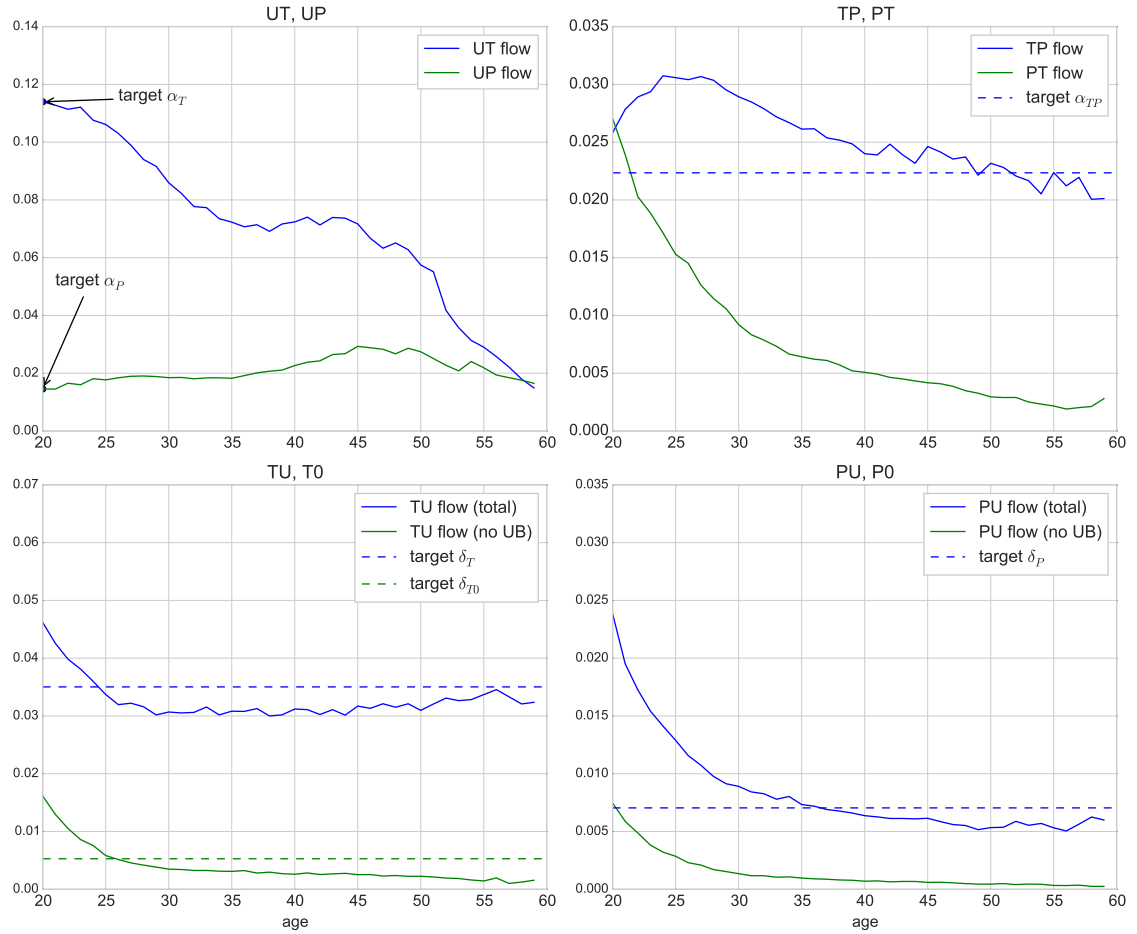
I assume a simple structure for search capital: three levels, that result in three proportional job finding rates ( $\alpha_j(s) = \bar{\alpha}_j s$ ). Search capital parameters are not pinned down, so they could be estimated by minimising the distance to some moments in the data. Instead I take estimates of  $s$  by targeting the differences in job finding rates at age 20 with different number of jobs held before. Using the assumption that workers enter the labour market with no assets, figure 3.8 shows on the right the implied duration profiles at the three different levels of search capital (so that  $s_2 > s_1 > s_0$ ) while the left panel shows a histogram of duration of unemployment in the data, separated by number of temporary contracts. The substantial differences at short durations of unemployment for even one temporary contract suggests that

---

<sup>20</sup> $\alpha_{PT}^1 = \alpha_{PT}^0 / S_{PT}$ , where  $S_{PT}$  is the proportion of permanent workers age 20 that accept an average temporary wage offer.



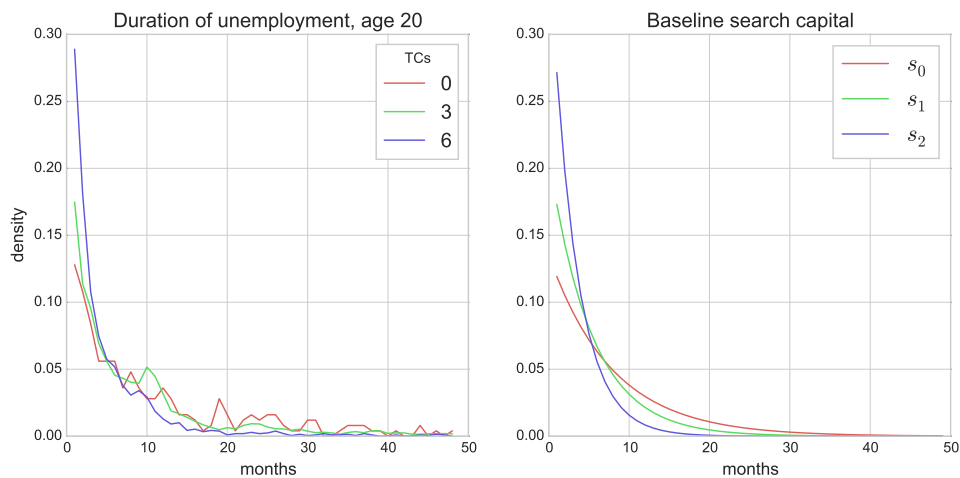
Figure 3.7: Monthly Transition Rates by Age (pre 2008, SS)



Source: Own calculations from MCVL, 2005-2013 waves

setting  $\pi_{s'|s}^+ = 1$  is reasonable. I repeat the exercise with smaller  $\pi_{s'|s}^+$  to match the differences between the shortest durations. For  $\pi_{s'|s}^-$ , the depreciation rates of search capital, I target a search capital loss every 5 years of tenure. Repeating the exercise with alternative depreciation rates (10,20,25 years) yields very similar results. In order to match the data the initial distribution of search capital is more important.

Figure 3.8: Duration of unemployment by number of contracts and search capital level



Source: Own calculations from MCVL, 2005-2013 waves

### Initial distributions

In order to match the job finding rates in the data it is important to acknowledge that some workers enter the labour market with a job in hand. I set the initial distribution of workers among states (unemployed with and without benefits, employed with a temporary and permanent contract) to match the data at age 20. For assets I assume all individuals start from 0 at age 20. For search capital, it is important to acknowledge that some workers aged 20 have had temporary jobs before entering

unemployment for the first time. Therefore I choose give the lowest level of search capital ( $s_0$ ) to new entrants (unemployed with no benefits) and level one to the rest. This is because in the data those receiving unemployment benefits must have accumulated enough job experience to be able to claim benefits. And indeed for unemployed workers less than 25 years old the average number of temporary jobs held before unemployment is lower among those without unemployment benefits (3 vs 5). Workers that enter the labour force with a job at hand are also assumed to be regular workers.

### 3.4.3 Results

#### Baseline Calibration

Figure 3.9 to 3.11 present the main results. Figure 3.9 shows the evolution of the shares of the three levels of search capital through workers lives, for all workers (left panel) and the unemployed (right panel). Over time, workers find stable jobs and search capital decreases: the proportion of bad searchers ( $s_0$ ) increases with time. During the first five years, the stock of good searchers ( $s_2$ ) increases as workers go through chains of temporary contracts, but then it decreases again to steady-state levels. By the time workers are 40, the overall search levels have reached steady state. Notice that the distribution is polarized by the end, with bad and good searchers having the highest shares. If we look at the unemployment pool, we see that at the beginning unemployed agents are not good searchers. Then as time goes by search capital increases - workers become more experienced in the labour market - but only at the end the stock of very good searchers increases as well. This is interesting because it shows that older unemployed workers are better searchers on average, so their lower transition rates to employment come from their higher levels of self

insurance that allows them to be more selective with their future jobs. This is the same mechanism as in the turbulence literature, which proves adequate to describe older worker dynamics.

Figure 3.9: Search Capital by age

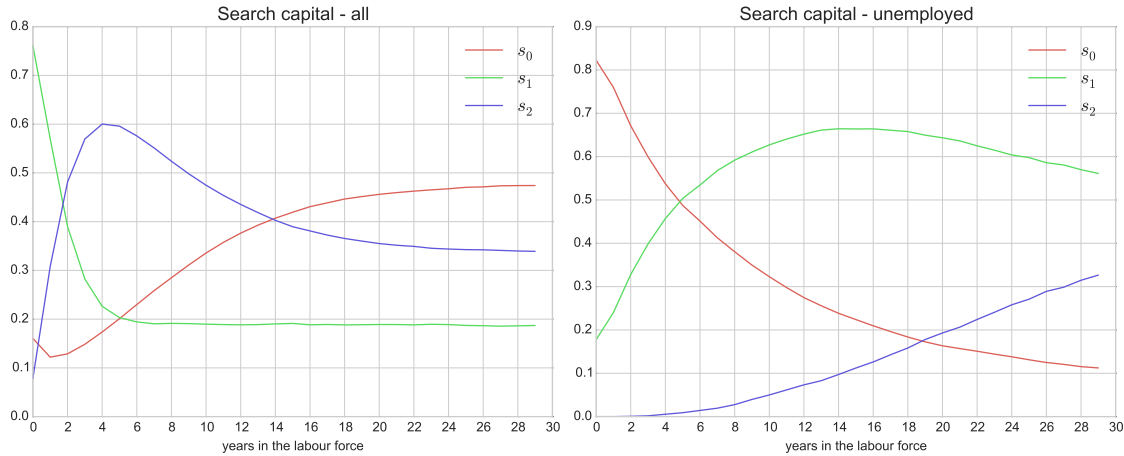
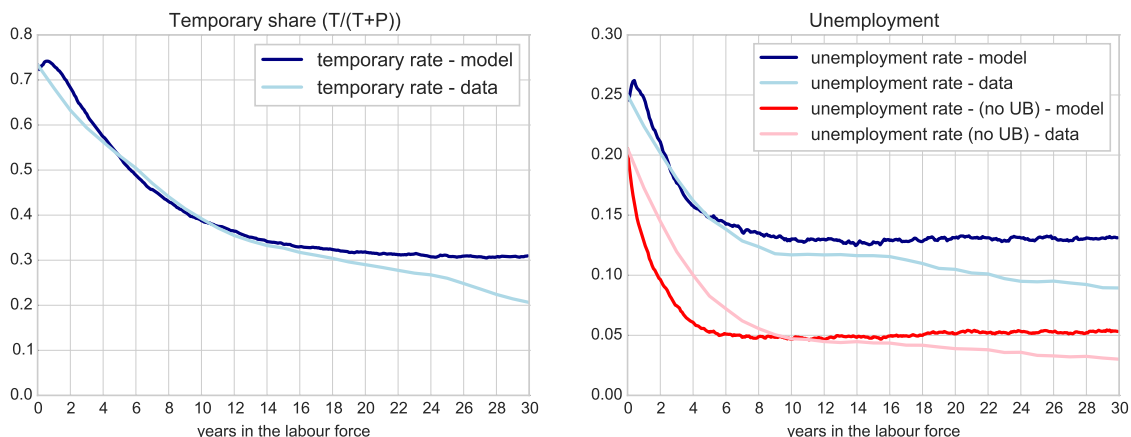


Figure 3.10 shows the evolution of unemployment and the temporary share of employment in the data and in the model. Having an accurate description of the temporary share of jobs is important for matching the average number of jobs workers have. The left panel of figure 3.10 shows that the model performs well, matching the temporary share for all ages until age 45, where it stays constant instead of steadily declining. These are the effects of the infinite horizon problem. In terms of unemployment, the right panel of figure 3.10 shows that the model does particularly well for overall unemployment, and underestimates unemployment without benefits for the young. This last result can be explained by the constant unemployment benefit expiration rate ( $U_0$ ) as panel 6 in figure 3.11 shows. While the inclusion of tenure related severance payments could be added in a similar manner to Dolado et al. (2016), making expiration of unemployment tenure related is more complicated, as

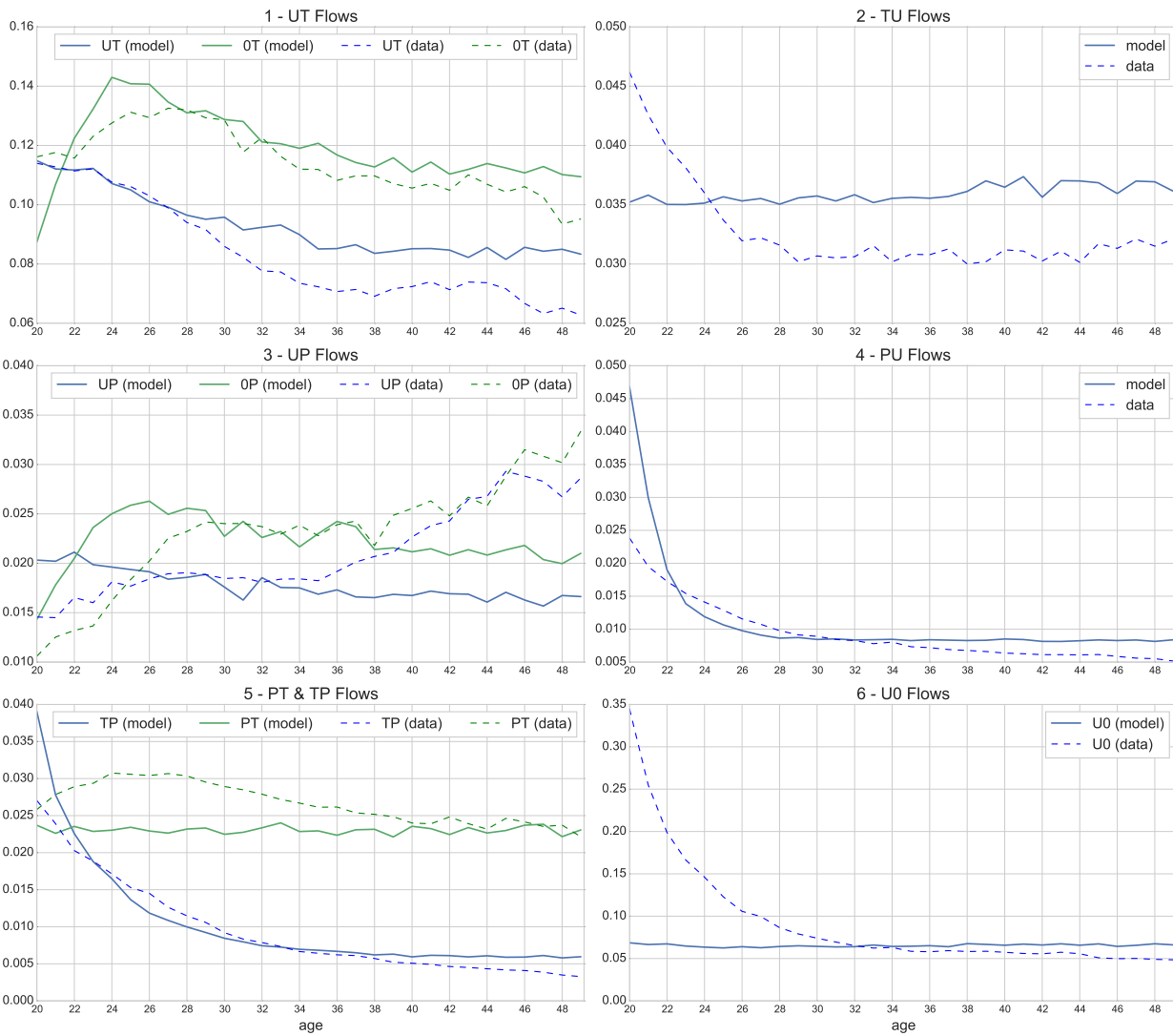
Figure 3.10: Temporary share and Unemployment rate



not only tenure but past tenure becomes a state variable. An alternative would be to use the framework developed by Andersen et al. (2017).

In terms of other flows, figure 3.11 shows yearly averages of monthly transition rates in the data and in the model, so that the data point at age 20 represents the 12 month average between years 20 and 21. The model matches the evolution of temporary job finding rates well (panel 1), with flows from unemployment without benefits (0T) being above those from unemployment (UT) both in the model and in the data. The evolution of search capital explains the high transition rates of those aged 20-30, and follows the decline over the 30s. Transition rates from regular unemployment are slightly above the data after age 30, staying above 8% per month, but they also follow a pattern of steady decline. Unemployed at age 30 are better searchers, but also have accumulated enough wealth and benefits to self insure against unemployment risk, allowing them to be more selective with the jobs they get. Flows to permanent jobs (UP, panel 2) are close to the data, but the model overestimates them for young people (less than 28 years old) and underestimates them for over 40's. This is likely to be what is driving the decline of the temporary share in the data (figure 3.10. A

Figure 3.11: Transition rates



likely explanation for this is the lack of transferable human capital: older workers may still carry some knowledge from their previous jobs that gets them better wages.

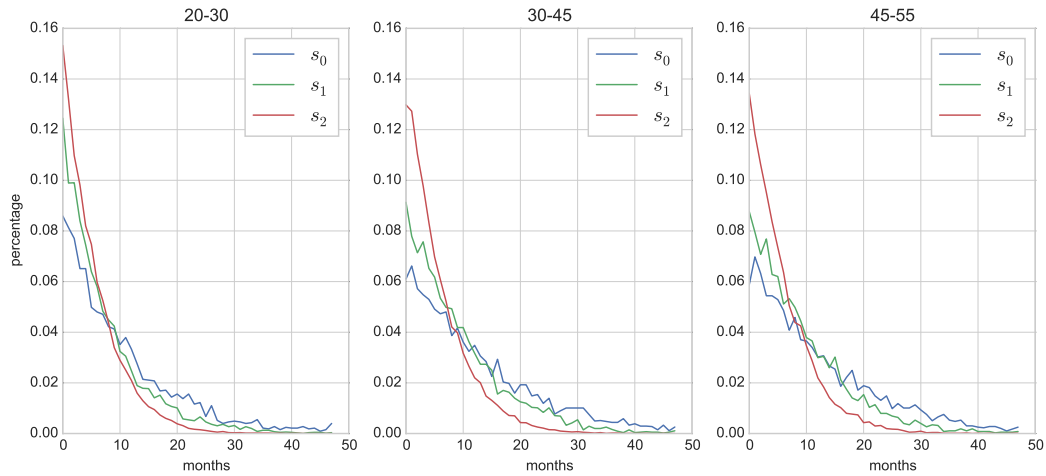
The permanent to temporary transition rate (PT, panel 5) follows its data counterpart closely: it overshoots at the beginning<sup>21</sup> but then follows the data remarkably well. These transitions defy conventional explanations, as in most models of dual labour markets with temporary contracts workers always prefer a permanent job to a temporary one. Here this is still true for permanent jobs with the same wage as a current job, but workers may be tempted to leave for better outcomes, even if these are uncertain. Switching jobs increases their search capital, which is another form of self-insurance against risks and allows them to climb the wage ladder. An alternative interpretation would be to increase the firing rate for permanent workers, but with advance notice - that is, giving workers time to find a temporary job before their current contract ends. This would add to the instability of permanent contracts for young workers.

Job destruction rates from temporary contracts (panel 2) do not follow the patterns in the data, which is to be expected since a constant job firing rate was assumed, however job separations from permanent contracts (panel 4) do follow the pattern in the data (overshooting at the very beginning). The difference here is that for young workers with a low wage permanent contract it is profitable to quit after some period of saving, to unemployment. Recall that the assumption, based on the data, was that quitters from permanent contracts get unemployment benefits, but that this is not the case for temporary jobs. This could be the driving difference, but there is an upwards trend in the temporary separation rate for older workers: older unemployed

---

<sup>21</sup>Recall that the flow plots are yearly averages: by construction the PT rate at age 20 matches the data exactly.

Figure 3.12: Duration of unemployment, by search capital



individuals may take temporary jobs to bump their savings up, and then quit. Their search capital is high on average (as figure 3.9 showed) and they can therefore afford to quit and search again.

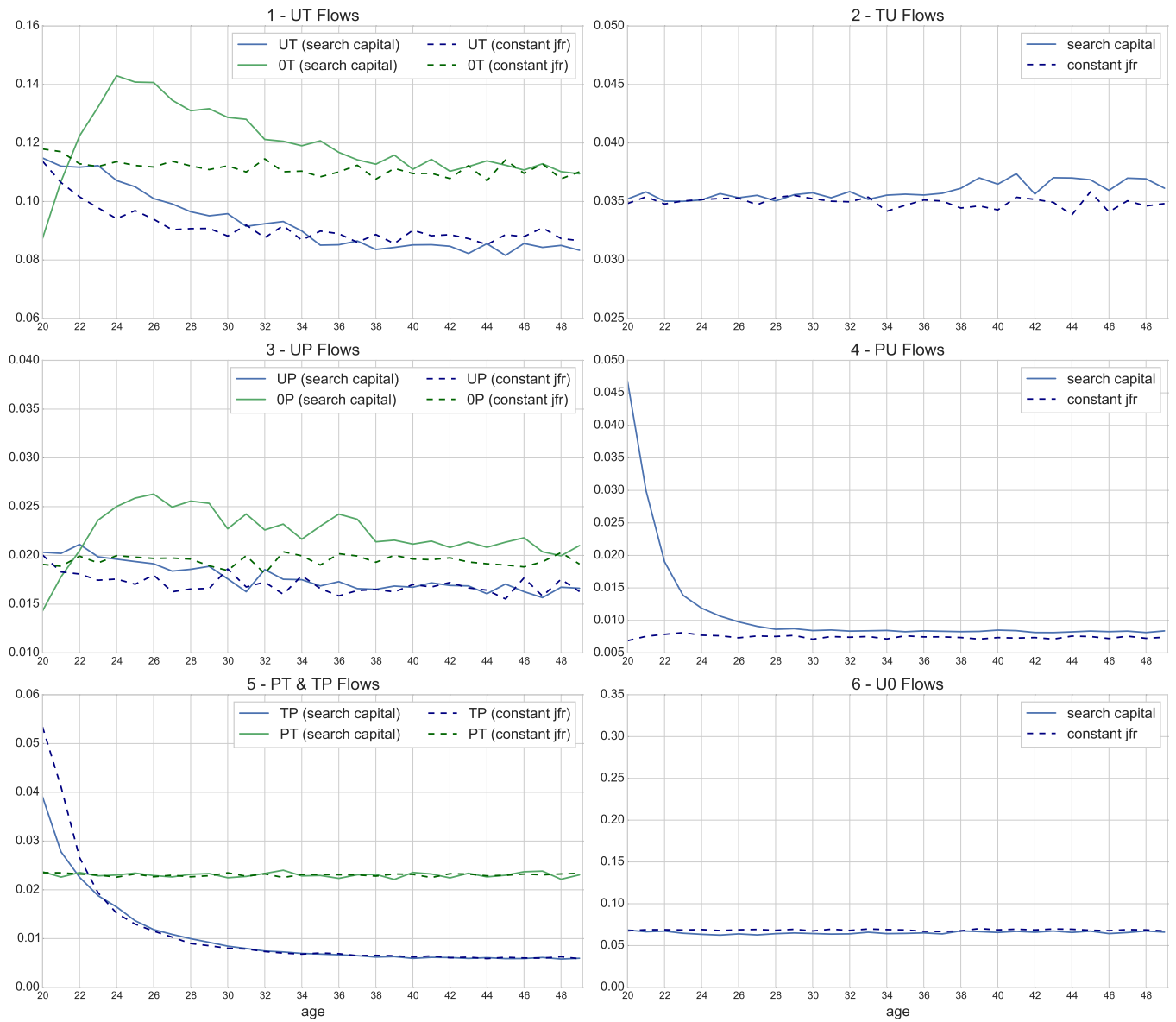
Finally, if we look at duration of unemployment by age, figure 3.12 shows that the gap between high skilled searchers ( $s_2$ ) and the rest widens with age, with low skilled searchers and moderately skilled searchers become closer. The bulk of these long term unemployed workers over 2 years is almost entirely made up of moderately and low skilled searchers ( $s_1$  and  $s_0$ ). Notice that if we considered all search abilities the duration distribution would become more skewed towards the extremes, which can cause observational duration dependence.

### No search capital

To better understand what search capital adds to the model, figure 3.13 shows monthly flows with and without search capital - so there is a constant job finding rate. The main difference is that job finding rates are flatter: the temporary



Figure 3.13: Transition rates without search capital



job finding rate among unemployed workers without benefits who are younger than 35 is lower and constant. For unemployed workers with UB the results are similar, indicating that its decline over time is driven by asset accumulation only. These patterns also hold for transitions to permanent jobs (panel 3). Another significant difference is the permanent separation rate in panel 4: under a constant job finding rate it is flat and low. Search capital can explain the patterns of the data better (as the same panel in figure 3.11 indicated) as young workers quit low wage permanent contracts in search of better outcomes. However here the absence of search capital makes these quits unprofitable. I leave further analysis for future work.

### **3.5 Conclusion**

Treating job search as a skill that can be gained and forgotten over time brings new insights to old problems. It provides an explanation as to why young workers can be stuck in long term unemployment, especially in a recession when they have to compete for fewer jobs with better searchers. For older workers, having outdated knowledge of the job market can also hurt their chances of finding one, even if they retain some of their past human capital. Labour markets in which some workers are over-protected from unemployment while others experience it very frequently exacerbate the differences in search capital, which could potentially expose the economy to sharp increases in long term unemployment.

Using a detailed administrative dataset I identify the number of temporary jobs held by a worker as a proxy for search capital, as temporary workers are more exposed to unemployment. Using tenure, work experience, wages in the last job and other controls, I regress duration of completed unemployment spells against the number

of temporary contracts held to date finding a significant negative correlation. The effects are still significant after introducing individual fixed effects. Years since last unemployment spell, a variable intended to capture search skill depreciation, is positively correlated with duration, while tenure and duration of unemployment benefits also seem to play a major role.

It could be that workers who are more exposed to temporary contracts find worse jobs, but regressions on future wages show a positive effect, both by reducing duration of unemployment (which is negatively linked to wages) and directly, although this last effect is more modest. The number of temporary jobs is negatively correlated with duration of the next job and probability of finding a permanent contract, but after controlling for fixed effects its coefficient turns positive. This suggests that as workers accumulate search experience they get better jobs, faster.

The empirical evidence provides support for search capital being significant for individual outcomes. To address the impact in the aggregate labour market I build a search model with savings and risk aversion and introduce search capital. I use the empirical wage distributions and transition rates for the years 2005-2008 in Spain to calibrate the model. The addition of search capital to the model helps to reconcile the patterns of unemployment and job finding rates through a worker's lifetime, especially for young people. As workers get older the unemployed get better at searching, but overall search capital decreases as workers find stable jobs. These results are based on a limited partial equilibrium setting, but a model with an aggregate matching function and firms could provide further insights, as could adding a more detailed human capital accumulation process and retirement. Search capital could enrich the hysteresis literature by improving the performance of models for

younger workers.

Finally, search capital adds a different perspective to the debate on labour market institutions and flexibility in Europe: more dynamic and flexible labour markets are more volatile but can also be more resilient to aggregate shocks. Active labour market policies can play a significant role in alleviating the negative effects of a segmented labour market.

# Bibliography

- Alonso-Borrego, C., J. Fernández-Villaverde, and J. E. Galdón-Sánchez (2005). Evaluating labor market reforms: a general equilibrium approach. Technical report, National Bureau of Economic Research.
- Alvarez, F. E., K. Borovickova, and R. Shimer (2015). A nonparametric variance decomposition using panel data. Technical report, Mimeo, University of Chicago.
- Andersen, T. M., C. Ellermann-Aarslev, et al. (2017). Job duration and history dependent unemployment insurance. Technical report, CEPR Discussion Papers.
- Bentolila, S., P. Cahuc, J. J. Dolado, and T. Le Barbanchon (2012). Two-tier labour markets in the great recession: France versus Spain. *The Economic Journal* 122(562).
- Blanchard, O. and A. Landier (2002). The perverse effects of partial labour market reform: fixed-term contracts in France. *The Economic Journal* 112(480).
- Bonhomme, S. and L. Hospido (2017). The cycle of earnings inequality: evidence from Spanish social security data. *The Economic Journal*.
- Card, D., R. Chetty, and A. Weber (2007). The spike at benefit exhaustion: Leaving the unemployment system or starting a new job? Technical report, National Bureau of Economic Research.

- Carrillo-Tudela, C. and E. Smith (2017). Search capital. *Review of Economic Dynamics* 23, 191–211.
- Carrillo-Tudela, C. and L. Visschers (2013). Unemployment and endogenous reallocation over the business cycle.
- Costain, J. S., J. F. Jimeno, and C. Thomas (2010). Employment fluctuations in a dual labor market.
- Couch, K. A. and D. W. Placzek (2010). Earnings losses of displaced workers revisited. *The American Economic Review* 100(1), 572–589.
- Cozzi, M. and G. Fella (2016). Job displacement risk and severance pay. *Journal of Monetary Economics* 84, 166–181.
- Dolado, J. J., E. Lalé, and N. Siassi (2016). From dual to unified employment protection: Transition and steady state.
- Dolado, J. J., S. Ortigueira, and R. Stucchi (2012). Does dual employment protection affect tfp? evidence from spanish manufacturing firms.
- García-Pérez, J. I. and F. Muñoz-Bullón (2011). Transitions into permanent employment in spain: An empirical analysis for young workers. *British Journal of Industrial Relations* 49(1), 103–143.
- Güell, M. (2003). Fixed-term contracts and the duration distribution of unemployment.
- Güell, M. and B. Petrongolo (2007). How binding are legal limits? transitions from temporary to permanent work in spain. *Labour Economics* 14(2), 153–183.

- Güell, M. and L. Hu (2006). Estimating the probability of leaving unemployment using uncompleted spells from repeated cross-section data. *Journal of Econometrics* 133(1), 307 – 341.
- Hornstein, A. (2012). Accounting for unemployment: the long and short of it.
- INE (2013). Encuesta de la población activa.
- INE (2017). Ocupados por tipo de jornada, sexo y rama de actividad. valores absolutos y porcentajes respecto del total de cada rama.
- Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. *The American economic review*, 685–709.
- Kitao, S., L. Ljungqvist, and T. J. Sargent (2017). A life-cycle model of trans-atlantic employment experiences. *Review of Economic Dynamics* 25, 320–349.
- Krueger, A. B. and A. Mueller (2010). Job search and unemployment insurance: New evidence from time use data. *Journal of Public Economics* 94(3), 298–307.
- Lalive, R. (2007). Unemployment benefits, unemployment duration, and post-unemployment jobs: A regression discontinuity approach. *The American economic review* 97(2), 108–112.
- Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of political economy* 117(5), 914–940.
- Ljungqvist, L. and T. J. Sargent (1998). The european unemployment dilemma. *Journal of political Economy* 106(3), 514–550.
- Ljungqvist, L. and T. J. Sargent (2008). Two questions about european unemployment. *Econometrica* 76(1), 1–29.

- Mortensen, D. T. (1970). Job search, the duration of unemployment, and the phillips curve. *The American Economic Review* 60(5), 847–862.
- OECD (2014). Education at a glance 2014.
- OECD (2017). Unemployment by duration: incidence.
- Silva, J. I. and J. Vázquez-Grenno (2013). The ins and outs of unemployment in a two-tier labor market. *Labour Economics* 24, 161–169.
- Stovicek, K., A. Turrini, et al. (2012). Benchmarking unemployment benefit systems. Technical report, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.
- Wadsworth, J. (1991). Unemployment benefits and search effort in the uk labour market. *Economica*, 17–34.







# Appendix A

## A.1 Step-by-Step guide to work with the MCVL

This appendix provides technical guidance on how to turn the raw csv files from the MCVL into a panel dataset in Stata, similar to the Labour Force Survey (LFS) panelling format. It extends Section 1.2.2 and includes more details on how to combine the different files, create and select new variables and build a panel. Do files that follow these steps can be provided upon request.

There are four main blocks: formatting, binding, implementing unemployment extensions and panel formatting.

### 1. Formatting

There are two ways of formatting the MCVL: year-by-year panel and retrospective panel:

- The **year-by-year panel** uses the information in all waves of the MCVL separately. This allows to take into account workers who are in some waves but not in others (see Garcia-Perez, 2009) and keeps the representatively of the population in every year.
- The **retrospective panel** uses information from the latest available year only. Although some representative of the sample is sacrificed, it is easier to study unemployment duration as there are no “cuts” or overlapping spells as in the year-by-year version.

As described in the main body of chapter 1, there may be applications where one or the other is preferable. In what follows I describe the year-by-year panel

approach, as it is more complex and appropriate for unemployment measurement. The same steps are needed for the retrospective panel, but there is no need to limit spell duration to be within the year they are reported, which makes things simpler.

## 1.1 Formatting Affiliation files

Open the ASCII files for each year, name the variables according to the MCVL guide. Be careful as the position of variables as it changes through years. Then format the dates of start (*alta*) and end (*baja*) of each spell, which together with the personal identifiers define each spell.

Next proceed to clean the overlapping spells (spells beginning before the end of the previous spell) as in Garcia-Perez (2008) cases a (total overlap) and b (partial overlap). For total overlaps, keep the longest continuous spell and drop smaller spells that happen at the same time. For partial overlaps, make the continuing spells start when the previous ends.

It is time to create the most important variable: labour market status during the spell. I chose to make this a string variable - which will be convenient when creating flows - but a labelled numerical value would work as well. For this we need to combine the information of 4 other variables:

- *Tipo de relacion con la seguridad social*, codes 700-800 correspond to unemployment benefit claimants. Mark as unemployed (“U”).
- *Tipo de contrato*, codes 400-900 correspond to temporary contracts. Mark as temporary, “T”.
- *Tipo de contrato*, codes 99-400 correspond to permanent contracts<sup>22</sup>. Mark as

---

<sup>22</sup>Note that some of these contracts may be *fijos discontinuos*, that is, permanent workers that only work for part of the year. They are different than temporary workers because they don’t have

permanent, “P”. There is an exception though: code 540 corresponds to partial retirement, so I mark this as out-of-the-labour-force (OLF). Also those whose variable *regimen de cotizacion* is 140 are in early retirement - so I mark them as OLF as well.

- *Regimen de cotizacion*, codes 500-600 correspond to self-employment. I mark them as “A” for Spanish *autonomos*<sup>23</sup>.

You may also want to create other auxiliary variables, such as an indicator for part-time contracts. For this you can refer to the accompanying .do file or refer to the official guide. The variable for this is *Tipo de contrato*.

**Important:** the variables *Empleador (forma Juridica) - Letra NIF de la Entidad Pagadora* and *Identificador (NIF/CIF) anonimado de la entidad pagadora* uniquely identify firms both in the affiliation files and in the fiscal file. If you want to use wage information, make sure to create a variable that joins both into one string variable. I call this variable *firmID* and move it right after the worker identifier.

## 1.2 Formatting pension files:

Name the variables according to the official guide. If using for labour market flows, most of this variables are irrelevant, but keep the dates (and format them accordingly) and the personal identification number.

---

a contract expiration rate and are protected by severance payments. If the reserchers wants to treat them differently, their contracts correspond to codes 300-400.

<sup>23</sup>If you want to be really precise, you should mark those whose *Regimen de cotizacion* is equal to 700-800 and 824-840 as self employed. These are the cases of farmers and sea captains.

### 1.3 Formatting personal information files:

Name the variables according to the official guide.

Special care should be taken with the variable *fecha de defuncion* that marks the death date of some workers who passed. The birth date should also be considered carefully as there are some likely mistakes - most famously a worker who was supposedly born in 1906 and was still working in 2005 - likely a coding error for 1960.

Ideally the 2011-2012 personal file should be the most up-to-date information as there was a census in 2011. It is recommend to impose the values of the education variables from 2012 onwards over earlier years, whenever possible.

There are some exceptional cases of repeated personal identifiers. I chose to keep the youngest of the two, but whatever criteria you use, keep only one so it can merge easily with the affiliation file.

## 2. Binding

### 2.1 Binding the files together: pensions

Year by year, open the formatted affiliation file and append the pension file to it. sort by date. If the final observation of a person is an entry from the pension file (easily identifiable because all affiliation file variable will be blank) then fill in their labour market status variable as OLF. Then fill in all the information missing in this last entry from the previous spell (place of residence, for example). Delete all other pension entries if you are only interested in labour market flows.<sup>24</sup>

---

<sup>24</sup>For example, Some disability or widowhood pensions can be received while in the labour force. If these are of no interest for the study, they can be committed.

## 2.2 Binding the files together: personal information

Merge the formatted personal file and the affiliation file (with pension information) together, using the personal identifier as joining variable. It should match virtually all cases. For the rare exceptions, I choose to keep the affiliation registered without personal information but drop the personal information entries without a matching affiliation entry.

Now it is a good time to drop all spells that happen before the current wave year - so 2006 only has spells active in the period 1st January 2006 onwards. This would ease the binding process below. Skip this step for the earliest year in the sample if you want keep retrospective information before the start of the sample<sup>25</sup>. You can keep retrospective information for all years, but that would make the merging years together process more cumbersome. Whenever in doubt, keep all spells.

**Important:** create a variable called *year* and set it equal to the file year. This will help you identify the information that each wave brings to the unified panel. Save the enriched affiliation files.

## 2.3 Binding all years together

Start with the earliest year in the dataset.<sup>26</sup> Drop all spells starting further than the 31st of December - modify the end date of the spell so it is 31st December. This right censoring should ensure the years are well matched together, so the 2006 file brings only spells active in 2006, the 2007 in 2007 etc. Append the next affiliation

---

<sup>25</sup>In practice this trimming will affect all observations that are active in later years, so if you are using many waves together keep all retrospective information and drop the repeated cases later, during the **Biding all years together** phase.

<sup>26</sup>It is recommended the 2005 wave as the absolute earliest to keep the sample representative of the population in that year. However you can use the 2005 file but name an earlier year as your starting point. However the further away from 2005, the worse the representativity of the sample.

file (with pension and personal information). Trim the spells over the end of the year and add the next file. Continue until you are left with the last wave. Do not censored this last year.

If you choose to keep retrospective information in each year, as you append each year erase duplicated spells. Here having created a variable for each year will come in handy, as it would help to identify identical spells but in different years (waves): they must share the same start date and the same *firmID* value. For the rare cases where firmID is not available (for example in unemployment spells) use the variable *Codigo de Cuenta de Cotizacion Principal (CCCP)* (right after the fiscal identifier variable) to identify duplicates spells.

At this point you should have one unique affiliation file with all waves joined together and no duplicated spells except from those that last beyond a calendar year - for example, a job that starts in May 2006 and ends in June 2009 should have 4 entries: one each for 2006,2007,2008 and 2009.

Depending on what you are interested in, it may be a good idea to create a *effective end date* variable that matches the latest end date for each spell - in the example above, set the effective end date as June 2009 in all entries. This way if you want to get statistics on tenure, you can either consider tenure up to the current year or total tenure in the sample. In the previous example, the first variable would be 8 months (May-December 2006) and the second 3 years (May 2006- June 2009). As a general rule it is better to create new variable than modify old ones, and this is particularly true with dates.



### 3. Extensions

#### 3.1 Contract modification adjustment

In many cases contracts change across the years - this is the case of temporary workers promoted to permanent contracts. The way these cases are recorded in the MCVL is not easy to deal with. Ideally, for the purpose of job market flows we would like to have separate entries for each kind of contract.

Look at the variable *Fecha de modificacion del tipo de contrato inicial o del coeficiente de tiempo parcial inicial*, towards the end of the affiliation file variables. If this variable is filled with a date, then there was a change in contract. The next variable, *Tipo de contrato inicial*, contains the original contract code of the job. Use the guide in step 1 to interpret its value.

Create an indicator variable that is equal to 1 if (1) the current wave year equals the year of the contract modification date AND (2) the type of contract is not the same as the original type of contract (for example, if there is a change from temporary to permanent (or vice-versa) or to part-time).<sup>27</sup>

Duplicate the spell in which the indicator variable is equal to 1. Change the type of contract of the first copy to be the original type of contract. Change the end date of this first copy to coincide with the modification date, and change the start date of the second copy to the modification date. Depending on how you want to treat tenure, you may want to extend this last change to the start date of all the other entries in posterior years to the contract modification. Now you have two spells for each job: one before the contract change and one after.

Repeat these steps with the variable *Fecha de modificacion del tipo de contrato segundo o del coeficiente de tiempo parcial segundo* and *Tipo de contrato segundo*.

---

<sup>27</sup>In case you are interested on recalls or temporary contract renewals, you may wish to also create a new entry even if the contract type doesn't change.

This is the second contract modification variable.

Be careful when recording the length of each spell before and after the contract change. In some applications you may be interested in the whole period (for example for tenure) but if you want to count temporary and permanent job experience separately you may want to treat the two contracts differently.

### 3.2 Unemployment Expansions

Before proceeding, sort all spells by worker id, labour market state and date (in this order). Number the spells in separate variables for each state - so for example, if a worker was unemployed in two separate periods, create a variable called *number of unemployment spell* (NoU) and set it equal to 1 for the first one and 2 for the second. Or if a worker had 9 temporary jobs, create a variable NoT and number them chronologically.

Sort again the sample by id and date of entry and exit. Fill in all the blanks in NoU equal to the previous NoU value and set 0 for all spells before the first unemployment value. Using this variable (NoU), create another variable counting the days the worker is employed at each year in between unemployment spells. This will give us the total number of days contributed to the social security, which we will use to calculate unemployment benefit entitlements.<sup>28</sup> Remember to reset this counter to zero each time there is a new unemployment spell.<sup>29</sup>

Create a variable equal to the end of each spell (call it *original ending*) that will be of use later to calculate the extension period.

---

<sup>28</sup>Self-employed workers do not contribute to the social security so **do not count self-employment spells**.

<sup>29</sup>Some workers can choose to “save” part of their unconsumed unemployment benefits for next unemployed period, in which case the time contributed by the next job won’t count towards the total. By resetting after each unemployment spell by default we make sure we only count the minimum possible time a worker could have contributed to the social security.

### *The LTU expansion*

First, join consecutive unemployment spells *within the year*: if both unemployment spells came from the same wave, and one starts immediately after the other, I consider them one single spell<sup>30</sup>. There are many cases of workers that received more than one subsidy at the same time (because of illness or family reasons) but they are part of the same unemployment spell.

Second, if there is a gap between the end of an unemployment spell and the beginning of the next job, extend the end date of the unemployment spell as to join the two. Make sure that the next spell is employment or self-employment, and not retirement. The reason for this is that we can't be sure that these workers are looking for a job - if they transition to retirement probably they were out of the labour force to start with. I choose to extend the spells of workers whose last entry is unemployment to the end of the sample<sup>31</sup>. This is crucial to account for all the workers whose benefits have expired and are still unemployed at the end of the sample. If your final year is beyond 2009, you should definitely do this as the number of unemployed workers without benefits reaches 50% in 2012.

Third, if the previous extension meant that the unemployment spell extended over the year of its original wave, duplicate the unemployment spell and set the wave year equal to the next year. If as a result the spell extends over two years, create two copies. For example, if a spell started and ended in 2009 but after the expansion it ends in 2010, create a duplicate of the original spell and modify its year so it belongs to 2010. This way there would be two copies: one in 2009 and one in 2010 - as it would be the case if unemployment benefits wouldn't have expired.

---

<sup>30</sup>Some authors want to make distinctions between unemployment benefits and unemployment subsidies - the latter referring to reduced amounts that some long term unemployed workers receive after running out of unemployment insurance. If so you may want to skip this step.

<sup>31</sup>I restrict this expansions to the cases when the end date of unemployment benefits is within the two years prior to the end of the sample.

### *The STU expansion*

In addition to the previous expansion, create a new unemployment spell if there is a gap between two jobs that lasts more than 15 days<sup>32</sup> and **at least one** of the following conditions are met:

1. the first job was self-employment
2. the first job ended in a quit (if the variable *Causa de Baja en Afiliacion* is 51)
3. by the end of the first job, the worker hasn't accumulated 12 months of continuous employment

In all of the previous conditions, the worker is not legally entitled to unemployment benefits, and thus we can interpret the period between jobs as unregistered unemployment.<sup>33</sup> You can further restrict these conditions by imposing that the firm identifiers of the two firms are different, so the worker is not being recalled to the same firm.

Set the end and start dates to fill in the gap between jobs. If this expansion takes the unemployment spell over the year of the wave, duplicate as in the previous case.

Finally, stata creates a new variable to identify copies and originals every time you duplicate observations. If your software of choice does not do that, make sure you have an indicator variable for these unemployment spells so you can identify them later.

---

<sup>32</sup>This threshold is arbitrary. Results do not change much when the limit is put at 10 days or 1 month. Garcia-Perez (2008) also sets 15 days as a reasonable threshold.

<sup>33</sup>See section 1.3.3 in Chapter 1 for more information on this assumption.

## **4. Panel Formatting**

### **4.1 Select the window**

The LFS runs interviews during the reference quarter, and so it gets its answers from replies in an unknown reference day within the quarter. This is inevitably going to lead to discrepancies in the results, as if the reference day in the MCVL doesn't coincide with the LFS the answers can be different. The extent of the discrepancy would depend on the frequency of flows: if there are more transitions within the month than within the quarter, then the probability of discrepancy is higher. The approach here is to select a window period within the quarter (or the month if interested monthly transitions). I chose the 15th to the 30th of the first month of each quarter. That is, the 15th-30th of January, April, July and October. Avoiding the first of these months is important, especially in the case of January as many jobs start after the Christmas break - which in Spain can last up until the 6th of January. Several robustness checks can be done by changing the window, but the results don't substantially change - except in the Christmas season as noted.

### **4.2 Create quarterly state variables**

Once the window has been chosen, I focus on the spells whose entry date is after the beginning of the window period, but before the end. If there is more than one such spell within a window, I chose to keep the one that continues onwards - that is, the last one. Another approach is to take the one with longest duration, but this can prove difficult. For example, take a long employment spell that ends the 22th of January, followed by 6 months of unemployment. The first spell would be selected if we apply the longest duration rule, but the second spell would be selected instead if we keep the continuing spell. This is particularly important if in the next

window period the worker is employed again, as not counting unemployment can understate the flows in and out of unemployment. Because of this I choose to follow the continuing spell approach.

Once we have one spell per window, all that is left is to fill in a variable for a spell-period. In many cases you will have to create copies of the same spell, when it appears in two different windows. For example, an employment spell that features in the first and second quarter would need to be duplicated. The original will be assigned to quarter 1 and the copy to quarter 2.

If you have applied any of the unemployment expansions, you may also want to create a different state for unregistered unemployment spells. For example, if an unemployment spell that originally lasted for a quarter now lasts for two - because of the LTU expansion - then you can label the first observation "U" and the second "0". For this you must use the extension date variable we created before modifying the start and end dates of spells. All unemployment spells generated from the gaps expansion can also be labelled differently.

### **4.3 Create stocks and flows**

The last step is to only keep spells that feature in a quarter and discard all other spells. We will be left with a panel dataset that mimics the structure of the LFS, with one observation per quarter. In this case however we will have more detailed information - precise tenure and experience data for example. This can be used to calculate unemployment stocks (with or without benefits) and temporary share of employment, for example.

To create flows, just link two consecutive quarters for the same worker. String variables are best for this, as you can simply add two strings to form a new variable. Here you may want to relabel "TUT" or "UTU" flows, conditioning on the duration

of unemployment (or the temporary contract). This can also be achieved by following “the longest spell” rule when choosing one observation per quarter. Note that this brings the MCVL flows closer to the LFS, but they are not classification errors - which is the reason these flows are modified when working with surveys (see Elsby et al. (2015) for more details).

## A.2 STU additions

Table A.1 provides summary statistics for the unemployment spells broken by modification. The first to notice is that most unemployment spells (37%) are not modified, but the number of modified spells is also substantial (30% for the LTU expansion and 32% for the STU). It is worth noting that not all of these spells appear in the unemployment rates pictured in figures 1.4 and 1.7, only those that correspond to the state of the worker in a given quarter. The statistics in table A.1 correspond to all spells. The first thing to notice is that unemployed workers with a STU expansion spell are younger and have less experience in all type of employment and unemployment. In particular, they have been unemployed an average of 0.7 years in their working lives. The spells are also shorter than those of the LTU expansion (205 versus 430) but the original spells are even shorter. This is because of two reasons: First, a requirement on STU spells is that they last longer than 15 days (which is not required for registered unemployment spells). Second, many registered unemployed workers with long spells run out of benefits, which means that long spells are in the LTU expansion category. Figure A.1 illustrates this point by showing the histogram of spell duration by extension.

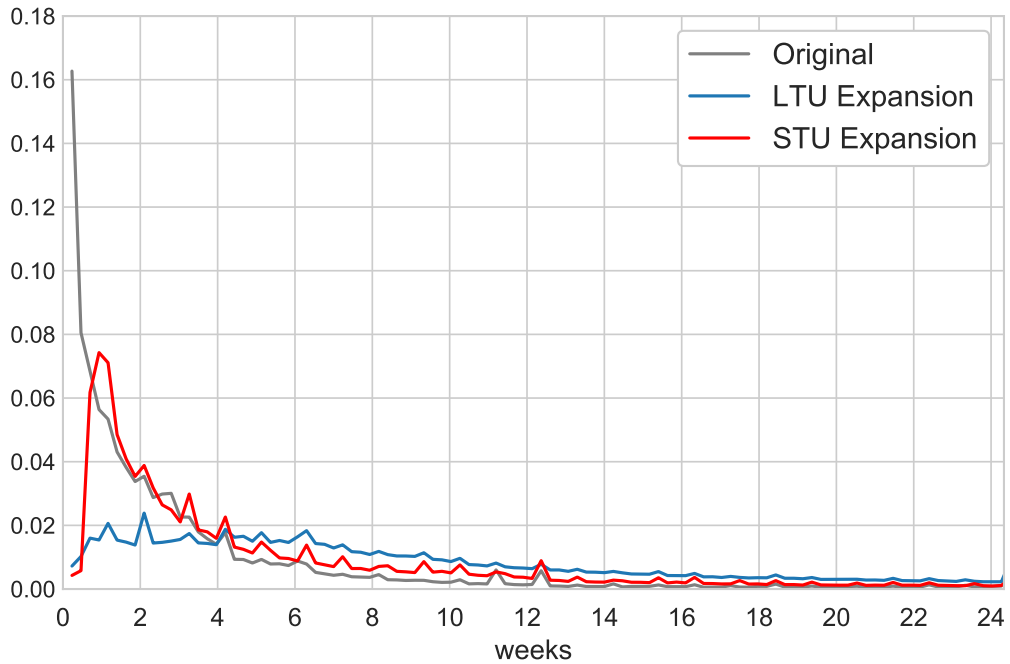


Table A.1: Summary Statistics of Unemployment Spells

	No Modification	LTU Expansion	STU Expansion
Observations	755,413	625,973	661,626
Percentage	0.370	0.306	0.324
Age	39.79 (10.79)	36.68 (10.15)	31.00 (10.12)
Female	0.462 (0.499)	0.445 (0.497)	0.449 (0.497)
Duration (days)	148 (322)	430 (494)	205 (286)
Experience PC	7.376 (8.502)	5.638 (7.108)	2.011 (3.928)
Experience TC	3.104 (2.633)	2.702 (2.311)	1.435 (1.653)
Experience Unemp	2.053 (2.227)	1.761 (2.003)	0.714 (1.349)

Sample is all unemployment spells in the 2004-2013 period. Averages with standard errors in parenthesis. Experience is measured in years,duration of the spell in days.

Figure A.1: Histogram of Spell duration, by Extension



## Appendix B: Detailed results for the Decomposition of Unemployment Duration

Table B.1: Variance decomposition with raw data and LTU expansion

	Spain raw data	Spain LTU exp.	Austria raw data
Total	1.409	1.694	1.711
Constant	1.645 (1.168)	1.645 (0.971)	1.645 (0.961)
DD	-0.527 (-0.374)	-0.341 (-0.201)	-0.187 (-0.109)
HT	0.290 (0.206)	0.390 (0.230)	0.253 (0.148)

Notes: Log duration in days, share of total variance in parenthesis. Ages 25 and over.  
Results for Austria taken from ?

Table B.2: Variance decomposition with LTU and STU expansion

	Spain NE	Spain STU exp.	Spain STU + spell corr.	Austria Non Employment
Total	3.240	2.607	3.257	3.081
Constant	1.645 (0.508)	1.645 (0.631)	1.645 (0.505)	1.645 (0.534)
DD	0.724 (0.223)	0.444 (0.170)	0.918 (0.282)	0.624 (0.203)
HT	0.871 (0.269)	0.519 (0.199)	0.694 (0.213)	0.812 (0.264)

Notes: Log duration in days, share of total variance in parenthesis. Ages 25 and over.  
Results for Austria taken from ?

Table B.3: Variance decomposition along the business cycle

	Spain 2002-2007	Spain 2008-2013
Total	2.945	2.742
Constant	1.645 (0.558)	1.645 (0.600)
DD	0.685 (0.233)	0.601 (0.219)
HT	0.615 (0.209)	0.495 (0.181)

Notes: Log duration in days, share of total variance in parenthesis. Ages 25-50.  
Spain, fully expanded data.

# Appendix C

## C.1 Robustness Checks

Table C1 below shows the coefficients for the controls on table 3.3. In particular, education, sex, dummy for foreign born, industry of previous employment, dummy for quit, part-time in the previous job.

Table C2 shows the results of adding a quadratic term for number of temporary contracts in the main regressions of tables 3.3 - 3.6. Nothing really significant changes, although the quadratic term is significant for regressions on duration of unemployment (columns 1-2) and duration of next job (column 4). In all regressions, the number of temporary contracts needed to turn the sign of the effect is over 100.

Table C3 presents the results of the regressions on unemployment duration (table 3.3 by industry of next job. The coefficient on number of temporary jobs (No. T) is significant and negative in all regressions. The results are consistent to those in table 3.3.

Table C.1: Regressions on Duration - controls

	Pooled OLS			Fixed Effects		
	(1) log(weeks)	(2) log(weeks)	(3) log(weeks)	(4) log(weeks)	(5) log(weeks)	(6) log(weeks)
quit	-0.185*** (0.00394)	-0.302*** (0.00398)	-0.297*** (0.00401)	-0.200*** (0.00466)	-0.261*** (0.00566)	-0.259*** (0.00568)
Construction (post 2008)	0.545*** (0.00926)	0.838*** (0.0112)	0.868*** (0.0116)	0.358*** (0.00818)	0.523*** (0.0105)	0.537*** (0.0109)
Construction (pre 2008)	0.226*** (0.0104)	0.510*** (0.0115)	0.538*** (0.0118)	0.0650*** (0.00909)	0.228*** (0.0116)	0.240*** (0.0119)
male	0.0185*** (0.00368)	0.0253*** (0.00390)	0.0238*** (0.00389)	-	-	-
High School	-0.0523*** (0.00419)	-0.0464*** (0.00456)	-0.0463*** (0.00455)	-	-	-
Baccalaureate	-0.0782*** (0.00558)	-0.0691*** (0.00542)	-0.0694*** (0.00541)	-	-	-
College	-0.0591*** (0.00622)	-0.0467*** (0.00628)	-0.0474*** (0.00627)	-	-	-
part-time	0.0394*** (0.00432)	-0.00223 (0.00407)	-0.0125** (0.00405)	-0.0534*** (0.00478)	-0.0773*** (0.00558)	-0.0816*** (0.00559)
foreign born	-0.0523*** (0.00545)	-0.0576*** (0.00594)	-0.0560*** (0.00594)	-	-	-
Extractive Ind.	0.455*** (0.0316)	0.756*** (0.0334)	0.782*** (0.0337)	0.286*** (0.0463)	0.464*** (0.0510)	0.481*** (0.0516)
Manufactures (primary)	0.444*** (0.0112)	0.731*** (0.0122)	0.760*** (0.0125)	0.287*** (0.0111)	0.447*** (0.0135)	0.461*** (0.0138)
Manufactures (machinery)	0.453*** (0.0105)	0.755*** (0.0119)	0.785*** (0.0123)	0.330*** (0.0105)	0.497*** (0.0127)	0.511*** (0.0130)
Manufactures (detail and repair)	0.459*** (0.0143)	0.753*** (0.0156)	0.782*** (0.0159)	0.321*** (0.0174)	0.478*** (0.0197)	0.487*** (0.0199)
Energy, gas, residual treatment	0.482*** (0.0177)	0.782*** (0.0186)	0.814*** (0.0189)	0.299*** (0.0196)	0.465*** (0.0220)	0.482*** (0.0223)
Retail and repairs	0.486*** (0.00983)	0.772*** (0.0108)	0.805*** (0.0112)	0.299*** (0.00790)	0.454*** (0.0101)	0.471*** (0.0105)
Transport and storage	0.408*** (0.0107)	0.702*** (0.0128)	0.735*** (0.0130)	0.278*** (0.0116)	0.427*** (0.0141)	0.447*** (0.0145)
Hospitality	0.429*** (0.00896)	0.709*** (0.0109)	0.739*** (0.0113)	0.289*** (0.00872)	0.439*** (0.0113)	0.455*** (0.0116)
Communication and IT	0.455*** (0.0120)	0.739*** (0.0136)	0.768*** (0.0139)	0.259*** (0.0140)	0.412*** (0.0159)	0.424*** (0.0163)
Financial	0.423*** (0.0173)	0.719*** (0.0188)	0.751*** (0.0191)	0.238*** (0.0225)	0.410*** (0.0251)	0.430*** (0.0253)
Real State	0.489*** (0.0190)	0.765*** (0.0207)	0.793*** (0.0211)	0.318*** (0.0250)	0.458*** (0.0287)	0.469*** (0.0291)
Professional services	0.465*** (0.0110)	0.733*** (0.0123)	0.762*** (0.0126)	0.261*** (0.0105)	0.411*** (0.0126)	0.426*** (0.0129)
Auxiliary services (cleaning, gardening, rental)	0.398*** (0.00935)	0.658*** (0.0109)	0.685*** (0.0112)	0.221*** (0.00779)	0.362*** (0.00979)	0.376*** (0.0101)
Public Administration	0.565*** (0.0122)	0.848*** (0.0123)	0.879*** (0.0127)	0.302*** (0.0105)	0.462*** (0.0124)	0.478*** (0.0127)
Education	0.509*** (0.0125)	0.789*** (0.0131)	0.823*** (0.0135)	0.247*** (0.0125)	0.396*** (0.0146)	0.414*** (0.0149)
Health and Social Services	0.412*** (0.0113)	0.699*** (0.0126)	0.729*** (0.0129)	0.325*** (0.0124)	0.477*** (0.0144)	0.494*** (0.0147)
Other Services	0.461*** (0.0102)	0.724*** (0.0120)	0.752*** (0.0123)	0.261*** (0.0100)	0.393*** (0.0122)	0.405*** (0.0125)
Observations	587222	465832	461369	587222	465832	461369
Adjusted $R^2$	0.546	0.559	0.561	0.462	0.457	0.458
AIC	1502925.7	1189482.3	1176423.7	1082258.0	840389.4	829077.3

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm), self-employed and spells shorter than 15 days.

Table C.2: Quadratic term for temporary contracts

	(1)	(2)	(3)	(4)
	log(weeks)	log(weeks)	log(next wage)	Duration of next
	(OLS)	(FE)	(OLS)	employment spell (years)
No. T	-0.040*** (0.0007)	-0.014*** (0.0020)	0.009*** (0.0004)	-0.024*** (0.0005)
No. T <sup>2</sup>	0.000*** (0.0000)	0.000*** (0.0000)	-0.000 (0.0000)	0.000*** (0.0000)
YEmp	0.002*** (0.0007)	0.001 (0.0013)	0.002*** (0.0005)	0.008*** (0.0021)
3 months claim	0.222*** (0.0042)	0.177*** (0.0056)	0.067*** (0.0044)	0.102*** (0.0063)
6 months claim	0.224*** (0.0049)	0.199*** (0.0075)	0.120*** (0.0047)	0.166*** (0.0086)
12 months claim	0.171*** (0.0073)	0.194*** (0.0126)	0.148*** (0.0060)	0.102*** (0.0152)
18 months claim	0.135*** (0.0102)	0.173*** (0.0195)	0.145*** (0.0078)	0.015 (0.0247)
24 months claim	0.040** (0.0132)	0.109*** (0.0270)	0.073*** (0.0096)	-0.101** (0.0343)
Last P	0.032*** (0.0045)	0.036*** (0.0061)	0.019*** (0.0043)	0.096*** (0.0073)
Tenure	0.015*** (0.0012)	0.026*** (0.0023)	0.000 (0.0008)	0.043*** (0.0039)
Experience	-0.007*** (0.0005)	0.040*** (0.0034)	0.011*** (0.0003)	-0.002*** (0.0006)
log(past wage)	-0.080*** (0.0013)	-0.043*** (0.0017)	0.126*** (0.0018)	0.045*** (0.0017)
age	-0.001 (0.0016)	-0.023*** (0.0054)	0.043*** (0.0014)	0.020*** (0.0019)
log(weeks)			-0.106*** (0.0016)	-0.008*** (0.0020)
log(UI)				0.001*** (0.0001)
Constant	1.214*** (0.2354)	0.648 (0.4052)	7.132*** (0.2263)	-0.279 (0.3046)
<b>Controls</b>				
Years	✓	✓	✓	✓
Industry	✓	✓	✓	✓
Occupation	✓	✓	✓	✓
Region	✓	✓	✓	✓
Observations	465832	465832	465832	357914
Adjusted <i>R</i>	0.560	0.458	0.146	0.134
<i>AIC</i>	1188835	840068.483	1284936.423	1067594.896

Robust standard errors in parentheses. Sample is all finished unemployment spells ending in employment, for workers aged 20-55. Excludes recalls (workers returning to the same firm), self-employed and spells shorter than 15 days.

Table C.3: Regressions on Unemployment Duration, by next job industry

	log(wages)																		
	Agriculture	Extractive Ind.	Manufactures (1)	Manufactures (2)	Manufactures (3)	Energy and gas	Construction	Retail and repairs	Transport	Healthcare	Communications	Financial	Real state	Professional	Auxiliary services	Public Admin	Education	Health and so.	Other services
No. T	-0.018*** (0.0019)	-0.011 (0.0125)	-0.031*** (0.0023)	-0.022*** (0.0018)	-0.019*** (0.0014)	-0.038*** (0.0043)	-0.019*** (0.0008)	-0.021*** (0.0011)	-0.028*** (0.0012)	-0.026*** (0.0010)	-0.025*** (0.0016)	-0.024*** (0.0022)	-0.018*** (0.0072)	-0.029*** (0.0016)	-0.034*** (0.0008)	-0.034*** (0.0017)	-0.032*** (0.0023)	-0.029*** (0.0013)	-0.031*** (0.0013)
Year	-0.002 (0.0045)	-0.005 (0.0124)	-0.001 (0.0029)	0.001 (0.0056)	0.002 (0.0052)	0.001 (0.0063)	-0.001 (0.0035)	0.000 (0.0017)	0.001 (0.0030)	0.000 (0.0027)	0.008 (0.0046)	0.010 (0.0075)	0.017* (0.0068)	0.002 (0.0095)	0.008*** (0.0018)	0.001 (0.0024)	0.006 (0.0046)	0.009** (0.0032)	0.006 (0.0033)
3 months claim	0.185*** (0.0276)	0.121 (0.1059)	0.187*** (0.0233)	0.220*** (0.0217)	0.271*** (0.0438)	0.239*** (0.0508)	0.185*** (0.0101)	0.188*** (0.0104)	0.228*** (0.0226)	0.211*** (0.0117)	0.225*** (0.0296)	0.251*** (0.0501)	0.174*** (0.0545)	0.208*** (0.0290)	0.264*** (0.0098)	0.264*** (0.0213)	0.194*** (0.0213)	0.238*** (0.0189)	0.245*** (0.0189)
6 months claim	0.206*** (0.0315)	0.169 (0.1458)	0.205*** (0.0282)	0.220*** (0.0246)	0.211*** (0.0389)	0.232*** (0.0322)	0.108*** (0.0120)	0.237*** (0.0118)	0.193*** (0.0254)	0.221*** (0.0150)	0.252*** (0.0262)	0.241*** (0.0463)	0.157* (0.0416)	0.262*** (0.0227)	0.261*** (0.0121)	0.188*** (0.0192)	0.195*** (0.0243)	0.235*** (0.0221)	0.245*** (0.0233)
12 months claim	0.191*** (0.0489)	0.275 (0.1673)	0.175*** (0.0385)	0.207*** (0.0336)	0.217*** (0.0622)	0.181* (0.0701)	0.138** (0.0181)	0.196*** (0.0169)	0.200*** (0.0335)	0.179*** (0.0233)	0.217*** (0.0451)	0.267*** (0.0824)	0.092 (0.0851)	0.227*** (0.0336)	0.193*** (0.0185)	0.184*** (0.0276)	0.130*** (0.0387)	0.163*** (0.0227)	0.202*** (0.0366)
18 months claim	0.181** (0.0597)	0.112 (0.2063)	0.189*** (0.0401)	0.178*** (0.0441)	0.140*** (0.0482)	0.116 (0.1129)	0.145*** (0.0286)	0.177*** (0.0233)	0.206*** (0.0473)	0.085 (0.0374)	0.183** (0.0606)	0.150 (0.1124)	0.161 (0.1147)	0.157*** (0.0472)	0.153*** (0.0255)	0.172** (0.0408)	0.037 (0.0335)	0.045 (0.0439)	0.138** (0.0301)
24 months claim	0.228** (0.0850)	0.121 (0.2806)	0.100 (0.0644)	0.143** (0.0538)	0.218* (0.1007)	0.188 (0.1131)	0.090** (0.0340)	0.212*** (0.0294)	0.148* (0.0614)	-0.071 (0.0477)	0.161* (0.0804)	-0.082 (0.1627)	-0.078 (0.1458)	0.075 (0.0613)	-0.002 (0.0339)	0.045 (0.0790)	-0.061 (0.0790)	-0.101 (0.0637)	0.069 (0.0064)
Last P	-0.027 (0.0322)	0.179 (0.1254)	0.029 (0.0229)	0.066** (0.0227)	0.028 (0.0483)	0.019 (0.0470)	0.015 (0.0124)	0.042** (0.0101)	0.069** (0.0212)	0.031** (0.0110)	0.021* (0.0257)	0.009 (0.0501)	-0.002 (0.0569)	0.046* (0.0189)	0.017 (0.0104)	0.019 (0.0183)	0.088** (0.0288)	0.028** (0.0197)	0.028** (0.0130)
Tenure	0.021** (0.0077)	-0.006 (0.0292)	0.008 (0.0062)	-0.004 (0.0045)	0.004 (0.0082)	0.009 (0.0109)	0.015** (0.0032)	0.004 (0.0026)	0.018** (0.0054)	0.032** (0.0047)	0.006 (0.0079)	0.025 (0.0129)	0.004 (0.0118)	0.010 (0.0068)	0.029*** (0.0031)	0.017** (0.0043)	0.020** (0.0068)	0.027** (0.0056)	0.015* (0.0046)
No. P	-0.039*** (0.0067)	-0.003 (0.0346)	-0.051*** (0.0031)	-0.011 (0.0089)	-0.024 (0.0185)	-0.022*** (0.0043)	-0.021** (0.0029)	-0.008*** (0.0043)	-0.040*** (0.0060)	-0.035** (0.0020)	-0.021* (0.0089)	-0.010 (0.0178)	-0.004 (0.0216)	-0.004*** (0.0084)	-0.023*** (0.0036)	-0.039*** (0.0083)	-0.023*** (0.0039)	-0.015 (0.0039)	-0.046*** (0.0039)
Experience	-0.008** (0.0018)	-0.007 (0.0113)	-0.002 (0.0018)	-0.011** (0.0018)	-0.007* (0.0041)	-0.005 (0.0033)	-0.011** (0.0008)	-0.008*** (0.0010)	-0.010*** (0.0018)	-0.009** (0.0011)	-0.004 (0.0029)	0.002 (0.0062)	0.003 (0.0043)	-0.000 (0.0021)	-0.008*** (0.0009)	-0.008*** (0.0013)	-0.009*** (0.0021)	-0.011** (0.0015)	-0.003 (0.0018)
log(post wages)	-0.041*** (0.0067)	-0.066 (0.0485)	-0.077*** (0.0070)	-0.088*** (0.0072)	-0.075*** (0.0148)	-0.072** (0.0130)	-0.065** (0.0030)	-0.085*** (0.0032)	-0.075** (0.0071)	-0.083** (0.0037)	-0.082** (0.0071)	-0.130*** (0.0130)	-0.042* (0.0175)	-0.091*** (0.0053)	-0.081*** (0.0027)	-0.064*** (0.0049)	-0.085*** (0.0036)	-0.128*** (0.0034)	-0.095*** (0.0031)
age	0.004 (0.0071)	0.044 (0.0399)	0.002 (0.0066)	-0.002 (0.0066)	0.009 (0.0123)	-0.024 (0.0135)	-0.001 (0.0031)	0.007* (0.0033)	-0.013 (0.0071)	-0.010* (0.0047)	-0.026** (0.0094)	0.015 (0.0190)	-0.005 (0.0167)	-0.008 (0.0066)	-0.003 (0.0039)	-0.004 (0.0071)	0.020** (0.0055)	0.009 (0.0055)	-0.010 (0.0059)
Constant	1.478*** (0.1341)	0.474 (0.9730)	1.643*** (0.1002)	1.570*** (0.1014)	1.533*** (0.1374)	1.929*** (0.1270)	1.500*** (0.0300)	1.529*** (0.0884)	1.531*** (0.1312)	1.617*** (0.1170)	1.509*** (0.2404)	1.899*** (0.8034)	1.695* (0.4065)	1.909*** (0.1303)	1.220*** (0.0800)	1.969** (0.1270)	1.478*** (0.1311)	0.606** (0.1359)	1.247*** (0.1740)
Year dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
industry dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
occupation dummies	11034	506	13055	16350	4323	3489	6725	64521	16371	51002	11134	3115	2222	19050	87080	23138	14389	28102	22949
Adjusted R <sup>2</sup>	0.575	0.484	0.544	0.510	0.574	0.586	0.399	0.517	0.574	0.592	0.370	0.566	0.506	0.565	0.565	0.617	0.576	0.600	0.595
FIC	25050	1548	37538	41103	10791.130	8886	162939	154828	43103	129222	28224	7806	5312	40479	224007	55236	39398	82515	55867

Robust standard errors in parentheses. Sample is all funded unemployment spells ending in employment for workers aged 20-55. Excludes "results" workers returning to the same firm, self-employed and spells shorter than 15 days.