The Cycle of Earnings Inequality: Evidence from Spanish Social Security Data*

Stéphane Bonhomme
University of Chicago
sbonhomme@uchicago.edu

Laura Hospido
Bank of Spain and IZA
laura.hospido@bde.es

This draft: June 9 2015

Abstract

We use detailed information from social security records to document the evolution of male earnings inequality and employment in Spain from 1988 to 2010. We find that inequality was strongly countercyclical: it increased around the 1993 recession, experienced a substantial decrease during the 1997-2007 expansion, and then a sharp increase during the recent recession. This evolution went in parallel with the cyclicality of employment in the lower-middle part of the wage distribution. Our findings highlight the importance of the housing boom and bust in this evolution, suggesting that demand shocks in the construction sector had large effects on aggregate labor market outcomes.

JEL classification: D31, J21, J31

Keywords: Earnings Inequality, Social Security data, Unemployment, Business cycle.

*We would like to thank Jorge de la Roca for his help with the data in early stages of this project. We also thank Samuel Bentolila, David Dorn, Cristina Fernández, Luis Garicano, Aitor Lacuesta, Gerard Llobet, Claudio Michelacci, Josep Pijoan-Mas, Diego Puga, Ernesto Villanueva, Ken Yamada, two anonymous referees, the editor of the paper and seminar participants at various places for useful comments. Support from the European Research Council/ ERC grant agreement no. 0263107 is gratefully acknowledged. All remaining errors are our own. The opinions and analyses are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Bank of Spain or the Eurosystem.

1 Introduction

The literature has amply described and analyzed the increasing evolution of earnings inequality in the United States and other Anglo-Saxon countries.[1] Recent studies have also documented a steep increase in German inequality (Dustmann *et al.*, 2009, Card *et al.*, 2013). However, the other largest European countries have been comparatively less studied.[2] In this paper we consider the case of Spain. We find that the recent period was characterized by strong cyclical variations in both employment and earnings inequality.

The recent Spanish experience offers an opportunity to study the evolution of earnings inequality and employment in a period of large cyclical variations. During the last two decades, Spain has shown high levels and volatility of employment relative to other OECD countries. The period was characterized by a long expansion between two severe recessions: the 1993 recession, and the "great recession" that started in 2008. Variations in employment over the cycle were substantial: from 25% in 1994 the unemployment rate fell to 8% in 2007, before increasing again to 21% in 2010.

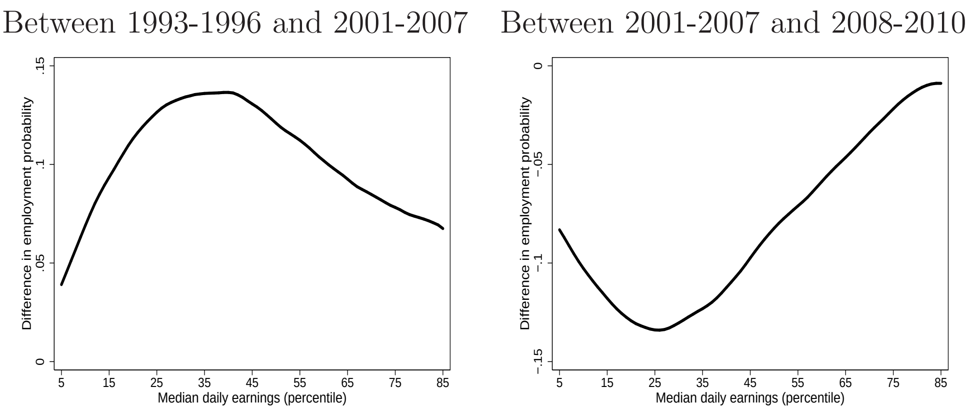
Using recently released social security data, we document that these employment fluctuations affected differently workers at different points of the earnings distribution. The left graph of Figure 1 shows the nonparametric regression curve, when regressing the difference between an individual's employment probability during the expansion and his employment probability around the 1993 recession (y-axis) on his rank in the distribution of median daily earnings during the period (x-axis). The right graph similarly compares the 2008 recession with the expansion. Throughout the paper we restrict the analysis to male workers because of data limitations. We see that both the employment gains during the expansion, and the losses during the recent recession, were larger in the lower-middle part of the distribution of daily earnings than in the tails. In Spain, the sensitivity to business cycle fluctuations has been highest for lower-middle wage workers.

A consequence of these employment fluctuations was to amplify the cyclical evolution of earnings inequality. Figure 2 shows the evolution of the logarithm of the 90/10 percentile ratio of male daily wages– a commonly used measure of inequality– between 1990 and 2010. Inequality closely followed the evolution of the unemployment rate: during

[1] Among the many references for the US see Bound and Johnson (1992), Katz and Murphy (1992), Levy and Murnane (1992), Acemoglu (2002), or more recently Autor *et al.* (2008). For the UK and Canada, see for example Gosling *et al.* (2000) and Boudarbat *et al.* (2006), respectively.

[2] For country-specific studies, see Manacorda (2004) on Italy, and the special issue of the *Review of Economic Dynamics* on Cross Sectional Facts for Macroeconomists (January 2010, 13(1)). Piketty and Saez (2006) provide a historical perspective for several OECD countries.

Figure 1: Employment growth as a function of daily earnings (males)

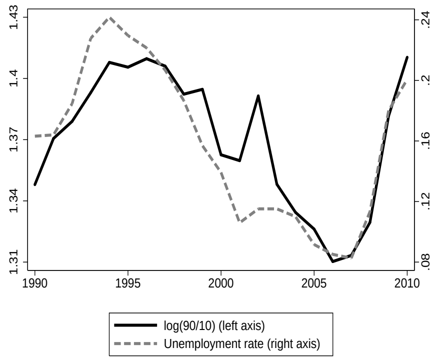


Notes: Source Social Security data. y-axis: difference in percentage of days worked by an individual relative to days present in the sample, between 1993-1996 and 2001-2007 (left), and between 2001-2007 and 2008-2010 (right). x-axis: rank of an individual in the distribution of median daily earnings during the period. Local linear regression, bandwidth chosen by leave-one-out cross-validation.

the 1997-2007 expansion inequality *decreased* by 10 log points, while between 2007 and 2010 it *increased* by the same amount. These are large fluctuations by international standards; for example in the US male inequality increased by 16 log points between 1989 and 2005 (Autor *et al.*, 2008). Comparing Figures 1 and 2 suggests a close link between the countercyclicality of inequality and changes in employment composition over the cycle, as inequality fell during the expansion when employment increased in the middle of the distribution, while it increased in the recent recession when a large share of lower-middle wage workers lost their jobs. Moreover, although we mostly focus on daily earnings inequality of employed workers, thus following a large literature on wage inequality that aims at separating wage effects from labor supply effects, we also document that monthly and annual earnings inequality experienced a similar countercyclical evolution. We further show that the inequality pattern remains qualitatively similar when using different approaches to include unemployed individuals in the calculations.

We consider several candidates to explain the joint evolution of employment and earnings inequality in the period. One particular factor is the recent evolution of the construction sector. Driven by the 1998-2007 housing boom, and then by the 2008 housing bust, employment in construction experienced a pronounced procyclical evolution, fluctuating between 13% and more than 20% of male employment. Construction-related sectors are also among the ones that experienced the strongest employment growth during the

Figure 2: Earnings inequality (males) and unemployment in Spain, 1990-2010



Notes: Source Social Security data and OECD. Logarithm of the estimated 90/10 percentile ratio of daily earnings of Spanish males (left axis), and unemployment rate (right axis).

expansion, and the steepest decline in the recent recession. Moreover, on average, construction workers belong to the lower-middle part, but not the left tail, of the earnings distribution. The effects of housing boom and bust on the labor market thus provide a possible explanation for the evidence pictured in Figures 1 and 2.

In order to quantify the importance of the construction channel, and more generally of changes in sectoral employment composition and changes in the wage structure ("price effects"), we perform various decomposition exercises. Specifically, we follow the methodology of Autor *et al.* (2005), and account for measures of skills (occupation and education groups), experience, and sectors. We find that both composition and price effects contributed to the decrease in inequality during the expansion. In contrast, when accounting for sectors in addition to skills and experience, composition changes fully explain the steep inequality increase in the 2007-2010 recession. This supports the idea that changes in employment composition, and in particular sectoral composition, have played an important role in the recent evolution of inequality. Moreover, when extending the analysis to account for other factors, we find that the large immigration inflow of the early 2000s, the duality of the Spanish labor market between permanent and temporary workers (Dolado *et al.,* 2002), and the evolution of the minimum wage, have had a relatively smaller influence.

To document these new facts on the Spanish labor market, our analysis relies on a recently released social security data set. In contrast with previous work based on cross-sectional and panel surveys, social security records have large sample sizes, wide coverage, and accurate earnings measurements. These data represent a unique source of consistent observations for a period of more than twenty years.[3] In a recent study, Dustmann *et al.* (2009) use social security data to provide an accurate description of the German earnings structure. Here we use individual earnings records to provide the first description of Spanish inequality over a long period of time.[4]

Although the social security data set is well-suited for the study of earnings inequality, it has two drawbacks. First, the data set has a proper longitudinal design from 2005 to 2010 only, whereas before 2004 the information is retrospective. This means that earnings data come from the records of individuals who were in the social security system some time between 2005 and 2010, either working, unemployed, or retired. Our comparison with other data sources suggests that, despite this retrospective design, past cross-sectional distributions of male (but not female) earnings remain representative up to the late 1980s. A second difficulty is that, as is commonly the case with administrative records, labor earnings are top and bottom-coded. To correct for censoring, we compare two approaches, and assess their accuracy using the tax files available in the most recent years for the same individuals as in the social security data set. Tax records are not subject to censoring, making them suitable to perform a validation check.

This paper finds a strong relationship between male daily and annual earnings inequality and the Spanish cycle. In the US, considerable attention has been devoted to explaining long-run trends in wage inequality abstracting from cyclical fluctuations, see for example Goldin and Katz (1998), Autor *et al.* (2003), and Lemieux (2008b).[5] Annual earnings inequality as measured by the 90/10 percentile ratio is also countercyclical in the US, as documented by Heathcote *et al.* (2010) and Guvenen *et al.* (2014) who study the cyclicality of individual earnings risk. In addition to the large magnitude of employment fluctuations, an important difference in the Spanish case is that the recent period has been characterized by a particularly long cycle by US standards– from the 1993 bust to the recent recession– close to twice the average length of US business cycles. This difference needs to be kept in mind when interpreting the results.

Our paper is also related to recent work on the cyclicality of employment in the

[3]In Spain, there is no other data set that reports information on labor income over such a long period. The longest running household survey is the Spanish labor force survey (EPA, in Spanish), which started in 1976. However, EPA does not contain any information on earnings.

[4]Felgueroso *et al.* (2010) use the same administrative source as we do, with the aim of documenting the driving forces behind the evolution of the earnings skill premium in Spain from 1988 to 2008. Ours is the first paper to use these data for the purpose of documenting earnings inequality.

[5]Barlevy and Tsiddon (2006) propose a model where secular changes in inequality are amplified in recessions.

US. Jaimovich and Siu (2013) find that middle-wage "routine" jobs disappear mostly in recessions. Although their definition of routine-manual jobs includes construction, they argue that the construction sector is not able to explain their findings. Purely cyclical factors are more likely explanations in the Spanish context, in particular because of a larger and more volatile construction sector than in the US. Charles *et al.* (2013) study the extent to which housing booms and busts, along with the secular decline of manufacturing, have determined the growth of US non-employment. Our findings suggest that, in the Spanish case, the interactions between the housing market and the labor market are also relevant to understand the evolution of aggregate employment and earnings inequality.[6]

Lastly, our description of the evolution of Spanish inequality is not inconsistent with previous work using survey data. In particular, like Pijoan-Mas and Sánchez-Marcos (2010), Carrasco *et al.* (2011), and Izquierdo and Lacuesta (2012) we find that earnings inequality decreased during the expansion period.[7] Compared to this literature, a longer-period view reveals the close link between earnings inequality and the business cycle, and the quality of the social security data allows us to conduct a precise quantitative analysis of changes in inequality.

The paper is organized as follows. As a motivation, in Section 2 we briefly discuss how changes in employment composition affect earnings inequality in a simple framework. We then describe the data in Section 3, and the evolution of earnings inequality in Section 4. Section 5 focuses on mechanisms. As a complement to the main analysis, in Section 6 we document the evolution of unemployment-adjusted measures of earnings inequality, obtained by imputing income values to the unemployed. Finally, Section 7 concludes.

2 Composition changes and inequality

In this section we outline the effect of a change in the composition of employment on earnings inequality, when the employment change affects the middle part of the earnings distribution. This situation characterizes the Spanish experience during the period where, partly driven by positive and negative demand shocks in the construction sector, employment fluctuations mostly affected lower-middle wage workers.

[6]Interestingly, recent papers provide evidence that the Spanish housing boom also had implications for education decisions (Aparicio, 2010, Lacuesta *et al.*, 2012).

[7]See also Farré and Vella (2008), Hidalgo (2008), and Simón (2009). Del Río and Ruiz-Castillo (2001), Abadie (1997), and Bover *et al.* (2002) provide evidence before 1990. Since the first version of this work was circulated, other papers have studied the recent evolution of inequality: Casado and Simón (2013) using the wage structure survey, and Bonhomme and Hospido (2013) and Arranz and García-Serrano (2013) using tax records.

Composition effects in a simple setup. To make the analysis simple and concrete, we consider an economy where changes in employment composition are driven by a demand shock in one particular sector ℓ. Employed workers, when working in sector j, earn a wage w_j. We focus on the impact on the earnings percentile ratio

$$R_\tau = \frac{F^{-1}(1-\tau)}{F^{-1}(\tau)},$$

where F is the aggregate cumulative distribution function (cdf) of wages, and τ is a percentage (typically, $\tau = 10\%$ or $\tau = 20\%$). R_τ is commonly interpreted as a measure of earnings dispersion or inequality.

The consequences of a sectoral demand shock in sector ℓ on earnings inequality depend on the relative position of ℓ in the earnings distribution. In the discussion, we abstract from within-sector differences in wages, and we assume that the wage in sector ℓ, w_ℓ, belongs to the middle part of the wage distribution in the sense that it lies strictly between the τ and $1-\tau$ wage percentiles, both before and after the demand shock.

As a result of the demand shock, employment in ℓ increases relative to other sectors. For simplicity, we assume that employment levels in other sectors $j \neq \ell$ evolve in the same proportion, and we abstract from the effect of the shock on sector-specific wages ("price effects"). In the supplementary appendix we present a simple equilibrium model with sectoral choice that has these features. Let δ denote the percentage change in the employment share of the sectors that are not directly affected by the demand shock. It can be shown that, due to the change in employment composition, the earnings percentile ratio becomes[8]

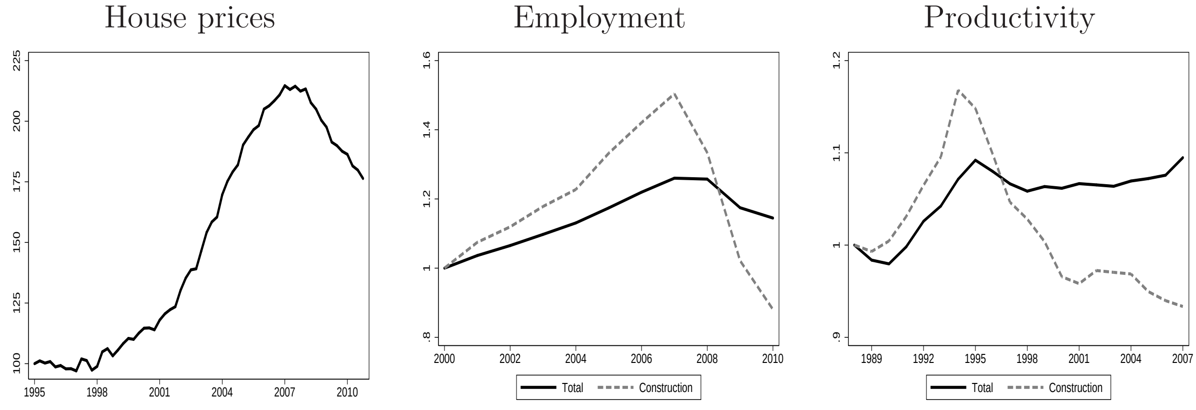
$$R'_\tau = \frac{F^{-1}\left(1-\frac{\tau}{1+\delta}\right)}{F^{-1}\left(\frac{\tau}{1+\delta}\right)}.$$

Hence, a positive demand shock in sector ℓ (which implies $\delta < 0$) leads to a *reduction* in earnings inequality. Intuitively, this decrease results from the fact that the middle part of the wage distribution grows relative to its tails. Similarly, a negative sectoral demand shock ($\delta > 0$) leads to an inequality *increase*. When applied to the Spanish case, this discussion highlights the relationship between the countercyclical evolution of inequality (documented in Figure 2), and the fact (documented in Figure 1) that employment fluctuations mostly affected workers in the lower-middle part of the wage distribution.

A candidate explanation: demand shocks in construction. In Spain, the housing boom and subsequent bust have contributed in an important part to employment

[8] See the supplementary appendix for a derivation.

Figure 3: House prices, employment, and productivity



Notes: Spanish ministry of housing and construction (left), Spanish national accounts (center), and EU Klems (right). Indices are normalized at the start of the period. Left graph: average real house price per square meter (quarterly). Center and right graph: solid line is total, dashed line is construction only.

fluctuations and changes in employment composition. Figure 3 provides three relevant facts. The left graph shows that real house prices per square meter more than doubled during the 1997-2007 housing boom. The causes of the boom are still a matter of debate, including low interest rates, the softening of lending standards in the mortgage market, the prevalence of homeowner tax deductions, large migration inflows, and the existence of overseas property buyers.[9]

The central graph in Figure 3 shows that, while total employment increased during the expansion and fell during the recent recession, employment in construction had a qualitatively similar but quantitatively much more pronounced evolution. Indeed, the fall between 2007 and 2010 amounts to nearly half of the population initially employed in that sector. As daily earnings of Spanish construction workers belong to the lower-middle part (but not the left tail) of the distribution, the above discussion suggests that these fluctuations may have played a role in the recent evolution of earnings inequality. Moreover, the effects of construction-driven composition changes are likely to be particularly large in Spain. As an example, employment in construction accounted for 11% of total employment (including males and females of all age groups) in 2000, compared to 5.8% in the US at the same date.[10]

[9]See, e.g., García-Montalvo (2007), Ayuso and Restoy (2007), González and Ortega (2009), Garriga (2010).

[10]Source: OECD. Variations in the employment share of construction were also lower in the US than in Spain, the share increasing to 6.3% in 2007 and decreasing to 5.4% in 2009. For non-college prime-age males, based on CPS data the construction share was 11% in 2000, 15% in 2007, and 11% in 2011 (Charles

Finally, the right graph in Figure 3 provides additional evidence of a demand shock affecting the construction sector. The graph shows the evolution of average labor productivity between 1988 and 2007, measured as value added per hours worked and computed from EU Klems data. While average productivity in the economy remained almost flat between 1995 and 2007,[11] productivity in the construction sector *fell* by 20% during the same period, consistently with a positive demand shock affecting that sector.

The empirical analysis below shows that changes in employment composition explain a substantial share of the evolution of Spanish inequality, particularly in the recent recession. It also highlights the special role of demand shocks in the construction sector. At the same time, our analysis of inequality takes into account a number of important factors that we have abstracted from in this section. In particular, we account for various dimensions of worker heterogeneity such as skills and experience, thus allowing for within-sector dispersion in earnings. The analysis also quantifies the empirical role of price effects, and accounts for the impact of labor market institutions (type of labor contract and minimum wage) and immigration. We now turn to the description of the social security data set.

3 The social security data set

3.1 Data and sample selection

Our main data source comes from the Continuous Sample of Working Histories (*Muestra Continua de Vidas Laborales*, MCVL, in Spanish). The MCVL is a micro-level data set built upon Spanish administrative records. It is a representative sample of the population registered with the social security administration in the reference year (so far, from 2004 to 2010). The MCVL also has a longitudinal design. From 2005 to 2010, an individual who is present in a wave and subsequently remains registered with the social security administration stays as a sample member. In addition, the sample is refreshed with new sample members so it remains representative of the population in each wave. Finally, the MCVL tries to reconstruct the labor market histories of the individuals in the sample back to 1967, earnings data being available since 1980.

The population of reference of the MCVL consists of individuals registered with the

et al., 2013). In our Spanish social security sample, the figures are 17%, 22%, and 14%, respectively. Educational achievement being under-estimated in the social security data, the Spanish figures are likely to be under-estimated as well.

[11] The slowdown of labor productivity growth between 1995 and the mid 2000s contrasts with the US and other European countries; see for example Dolado *et al.* (2011).

social security administration at any time in the reference year.[12] The raw data represent a 4 per cent non-stratified random sample of this reference population, and consist of nearly 1.1 million individuals each year. We use data from a 10 per cent random sample of the 2005-2010 MCVL.[13] To ensure that we only consider income from wage sources, we exclude all individuals enrolled in the self-employment regime. We keep prime-age male employees (aged 25-54) enrolled in the general regime.[14] Then, we reconstruct the market labor histories of the individuals in the sample back to 1980. Finally, we obtain a panel of 52,878 individuals and more than 7 million monthly observations for 1988-2010.[15] We present descriptive statistics on sample composition and demographics in Table A1 of the appendix.

The MCVL represents a unique source of consistent data for a period of more than twenty years. However, given its particular sampling design, using the retrospective information for the study of population aggregates may be problematic in terms of representativeness. In the supplementary appendix we consider three issues in turn. Mortality rates are too small to significantly affect the study of earnings inequality in the 25-54 age range. We also present evidence that attrition due to migration out of the country is unlikely to affect the results. In contrast, the evidence reported suggests that an important source of attrition for women is due to career interruptions, particularly in their 20s and early 30s. This is the main reason why we focus on males in the analysis.

Given that we rely on social security data, one might be concerned that the sectoral information could be incomplete, particularly for the construction sector where self-employment (that we remove from our main sample) and informal work might be prevalent. To show that this concern is unlikely to substantially affect our main results, in the supplementary appendix we compare the evolution of the employment share of construction in the MCVL and the Spanish labor force survey, and find similar numbers. We also show that the evolution of the employment share of the construction sector is similar when adding self-employed individuals back to our sample.

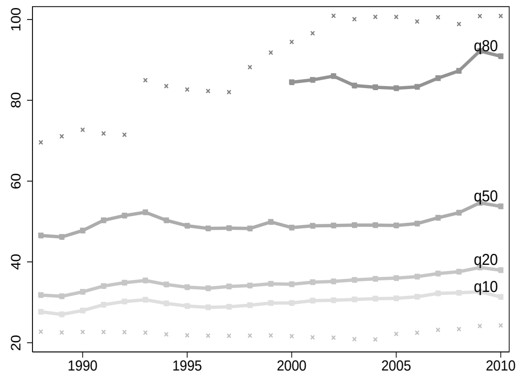
[12]This includes pension earners, recipients of unemployment benefits, employees and self-employed workers, but excluding individuals registered only as medical care recipients, or those with a different social assistance system (part of the public sector, such as the armed forces or the judicial power).

[13]This selection was done in order to reduce the size of the data set and ease the computational burden. Taking another 10% random sample made almost no difference to the results.

[14]In Spain, more than 95 per cent of employees are enrolled in the general scheme of the Social Security Administration. Separate schemes exist for some civil servants, which are not included in this study.

[15]The reason for starting in 1988 instead of 1980 is that sample representativeness tends to become less accurate as one goes back in time, as we document in the supplementary appendix.

Figure 4: Quantiles of uncensored daily earnings



Notes: Source Social Security data. Solid lines are observed quantiles of male daily earnings. Dark and light crosses are the real value of the maximum and minimum caps, respectively. Caps are calculated as averages of the legal caps over skill groups, weighted using the relative shares of each group every year.

3.2 Social security earnings and censoring correction

The MCVL provides information on the "contribution base", which captures monthly labor earnings plus 1/12 of year bonuses for those who work.[16] As is often the case in administrative sources, earnings are top and bottom-coded. The maximum and minimum caps vary over time and by occupation groups. They are adjusted each year with the evolution of the minimum wage and the inflation rate.[17] In most of the analysis, we use *daily wages* as our main earnings measure, computed as the ratio between the monthly contribution base and the number of days worked in that particular month. A daily wage measure is available for all workers who have been working at some point during the corresponding year. The social security data do not record hours of work, so we cannot compute an hourly wage measure.[18] Earnings are deflated using the 2006 general price index.

Figure 4 shows, for each year from 1988 to 2010, several percentiles of real daily earnings of Spanish males. The crosses on the graph represent the real value of the maximum and minimum caps. Real earnings have generally increased over the period. For example, median daily earnings increased from 46.5 Euros in 1988 to 54 in 2010.

[16] Exceptions include extra hours, travel and other expenses, and death or dismissal compensations.

[17] See the supplementary appendix (Figure S6). The groups are defined as follows. Group 1: Engineers, College. Group 2: Technicians. Group 3: Administrative managers. Group 4: Assistants. Groups 5-7: Administrative workers. Groups 8-10: Manual workers.

[18] The data contain measures of part-time and full-time work. Re-weighting daily earnings using these measures makes little difference (for males).

However, the proportion of top-coded observations is substantial: the 80th percentile is observed from 2000 to 2010, and the 90th is never observed. Hence, the 90/10 ratio is censored during the whole period. At the same time, the 50/10 ratio is never censored.

Censoring correction. We use a cell-specific tobit model to impute earnings to individuals whose earnings are censored (10 imputations per censored observation). In the baseline specification, covariates cells are based on occupation and age groups– our proxies for skills and experience, respectively– and yearly and monthly indicators (see the supplementary appendix). To assess the performance of this method, we take advantage of the fact that from 2004 to 2010 the MCVL was matched to individual income tax data, which are not subject to censoring. In the supplementary appendix we show that annual social security contributions and annual labor income obtained from the tax data are strongly correlated, although they are not identical. We find that the tobit method provides a good fit between the 10 and 90 earnings percentiles, clearly superior to the fit obtained using a linear quantile censoring correction method.[19] When interpreting the results, it will be important to keep in mind that the censoring correction is not perfect. Although comparison with the tax data suggests that it does a good job for the more recent period, the accuracy of the extrapolation may be poorer in the first part of the sample, where the amount of censoring is larger. In order to alleviate concerns related to the extrapolation, we will document the evolution of the 20th and 80th percentiles as a complement to the more commonly used 10th and 90th percentiles.

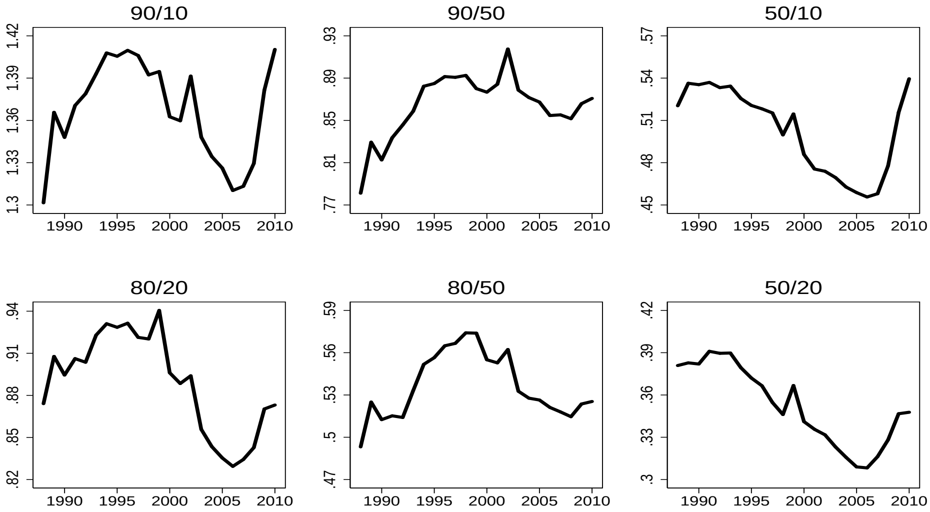
4 The evolution of earnings inequality

In this section we describe the evolution of male earnings inequality from 1988 to 2010. The top panel in Figure 5 shows the evolution of several inequality measures over the period: the ratio of the 90th to 10th daily earnings percentiles (90/10), the ratio of the 90th to 50th (90/50), and the ratio of the 50th to 10th (50/10), each of them in logs. Table A2 in the appendix reports the numerical values of the quantiles and the corresponding earnings percentile ratios for selected years. In addition, in the supplementary appendix we show the 10, 20, 50, 80 and 90th daily earnings percentiles (Figure S10).

Over the whole period, male inequality was markedly countercyclical. The 90/10 earnings ratio increased by 10.8 log points between 1988 and 1996, then decreased by 9.6

[19] In the supplementary appendix we also compare our results with the recent evolution of earnings inequality according to the tax data.

Figure 5: Log-percentile ratios



Notes: Source Social Security data. Log-ratios of estimated unconditional quantiles of daily earnings.

log points between 1997 and 2006, after which inequality increased again by 9.7 log points. In addition, the increase in male inequality during the earlier period was concentrated in the upper part of the earnings distribution, as the 90/50 earnings ratio increased by 11 log points, while the 50/10 earnings ratio remained stable. In contrast, the inequality decrease between 1997 and 2006 affected the two halves of the distribution, as the 50/10 ratio decreased by 6 log points, while the 90/50 ratio decreased by 3.6 log points. Moreover, the inequality increase in the recent recession mostly affected the bottom half of the distribution, with a 8.2 increase in the 50/10 ratio while the 90/50 ratio increased by 1.6 points only.

One concern with the 90/10 ratio is that it is sensitive to the censoring correction method. Given this, in the bottom panel of Figure 5 we show the 80/20, 80/50, and 50/20 percentile earnings ratios (in logs), which are less subject to censoring. The picture of male inequality is similar to the top panel, with a marked countercyclical pattern. Quantitatively, the changes are of a smaller magnitude, especially in the recent recession. For example, the 80/20 ratio increased by 5.7 log points between 1988 and 1996, decreased by 9.2 log points between 1997 and 2006, and increased by 3.9 between 2007 and 2010.[20]

[20]We also performed a number of robustness checks. As a first check, we re-weighted the data using mortality rates by gender and age groups, finding very similar results. As a second check, we re-weighted the monthly observations of daily wages in inverse proportion to the number of months worked in a year. The results are shown in the supplementary appendix (Figure S11). In that specification, inequality levels are higher than in the benchmark one, and the evolution is quite similar. The main differences appear

Table 1: Changes in log-percentile ratios, males (×100)

United States*		Spain**			Germany***	
1973-1989	1989-2005	1988-1996	1997-2006	2007-2010	1980-1990	1990-2000
90/10		90/10			85/15	
18.3	16.4	10.8	-9.6	9.7	8.3	10.7
90/50		90/50			85/50	
10.2	14.2	11.0	-3.6	1.6	5.8	5.1
50/10		50/10			50/15	
8.1	2.1	-0.2	-6.0	8.2	2.5	5.6

Notes: * Hourly inequality measures from Autor *et al.* (2008). ** Daily inequality measures estimated from Spanish Social Security data. *** Daily inequality measures from Dustmann *et al.* (2009)

A countercyclical evolution of inequality is also present when using other earnings measures, in addition to the daily measure that we consider here. In the supplementary appendix we show the evolution of log-percentile ratios of monthly and annual earnings instead of daily wages (Figures S12 and S13, respectively). In both cases inequality shows a strong decrease during the expansion, and a marked increase in the recent recession. Moreover, in Section 6 we document the patterns of inequality measures that include the unemployed in the analysis, by contrasting two different methods. All these different measures yield a qualitatively similar countercyclical evolution.

The fluctuations of Spanish inequality are substantial by international standards. To see this, consider the well documented case of the United States. According to Autor *et al.* (2008), and as reproduced in Table 1, male inequality measured by the 90/10 log-percentile ratio of hourly wages increased by 18 log points between 1973 and 1989. This corresponds to a yearly increase of 1%. A slightly lower yearly rate of increase in daily-earnings inequality was found by Dustmann *et al.* (2009) for Germany. In comparison, in Spain between 1997 and 2006 the 90/10 ratio *decreased* at a 1% rate per year, while between 2007 and 2010 it *increased* at a 2.4% rate per year.

The evidence presented in this section, while not inconsistent with previous work on earnings inequality in Spain,[21] offers two main novel descriptive insights. First, a longer-

during the recent recession: as a result of the higher weights given to the (mostly low-wage) individuals who work few months, the increase in the 90/10 ratio is larger in this alternative specification: 15 log points between 2007 and 2010. As an additional check, we re-estimated the percentile ratios focusing on workers with non-zero monthly earnings in all months within a year, finding results very similar to Figure 5.

[21] In the supplementary appendix we compare our results with recent papers that have attempted to document the evolution of Spanish inequality using other data sources.

period view shows that male inequality experienced a marked countercyclical pattern, the expansion period of fall in inequality being surrounded by two recession episodes where inequality increased sharply. Second, the quality of the social security data allows to document the quantitative magnitudes of these changes, which we find to be large by international standards.

5 Explaining the evidence

Here we document the impact of various factors on the evolution of earnings inequality and employment. We particularly emphasize the role of individual and employment characteristics (skills, experience, and sectors), while also accounting for labor market institutions (the minimum wage and the type of labor contracts) and immigration.

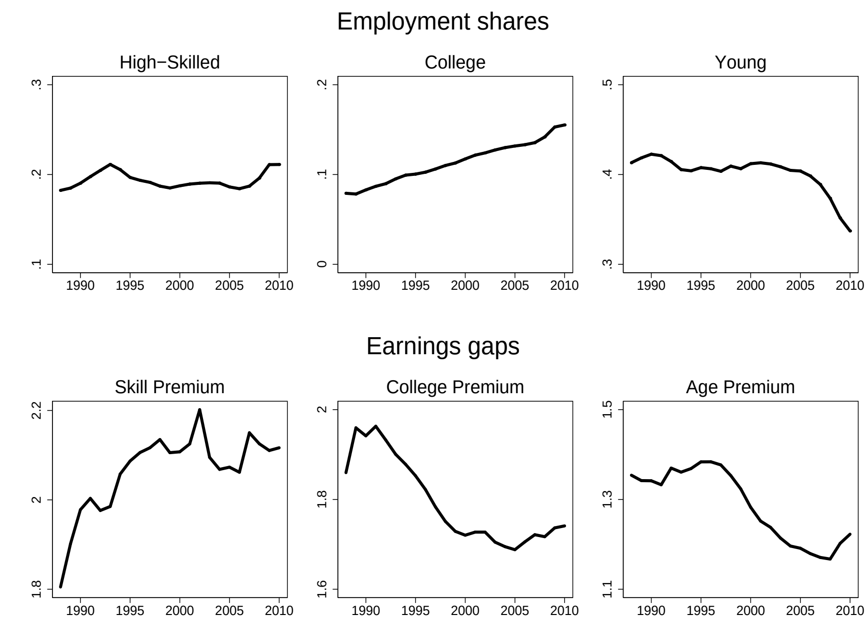
5.1 Skills, experience and sectors

We start by providing evidence on employment and earnings for different skill groups, experience groups, and sectors. This will help interpret the results of the decomposition exercises in the next subsection.

Skills and experience. The top graphs in Figure 6 show the employment shares of occupation groups (high versus low skills), education groups (college versus non-college), and age groups (younger than 35 versus older). The bottom graphs show median daily earnings gaps between these groups.

The bottom left graph in Figure 6 shows that the ratio of median daily earnings between high-skilled (occupation groups 1-3) and medium and low-skilled workers (groups 4-10) increased during the early 1990s, and remained approximately stable from 1997 to 2010. The central graph shows the evolution of the "college premium"; that is, the ratio between the median daily earnings of college graduates and those of non-college graduates. We see that the college premium decreased substantially from the early 1990s until 2005, by roughly 13%. This evidence of a decline in the college premium in Spain has been documented before (e.g, Pijoan-Mas and Sánchez-Marcos, 2010, Felgueroso *et al.*, 2010). We will see below that it has partly contributed to the fall in inequality during the Spanish expansion. Note also a slight increase in the college premium since 2005. The different evolution of the occupation and college earnings premia may in part be due to the fact that, as we see on the top graphs, the employment share of college graduates increased during the period, while the share of high-occupation groups remained relatively constant

Figure 6: Occupation, education, and age groups: employment shares and earnings gaps



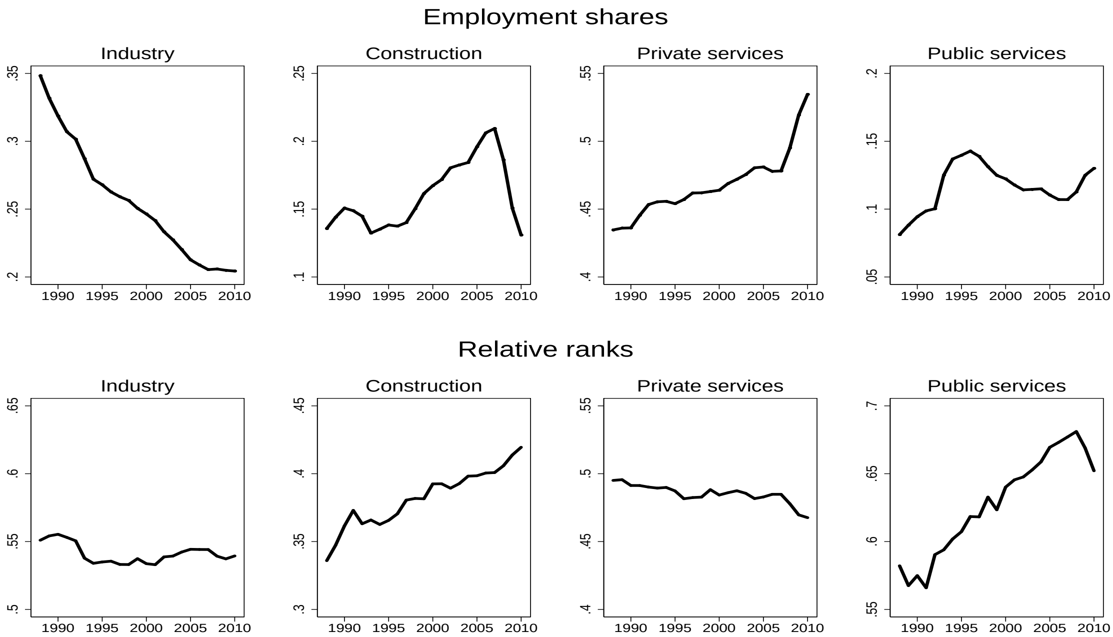
Notes: Source Social Security data. The top panel shows employment shares for various groups. The "premia" on the bottom panel refer to ratios of median daily earnings between i) occupation groups 1-3 and groups 4-10 ("skill premium"), ii) college and non-college workers ("college premium"), and iii) workers aged 35 years or more and those aged less than 35 ("age premium").

(except at the end of the period). Finally, the bottom right graph in Figure 6 shows the ratio of median daily earnings of older workers (35 years or more) and young workers. We observe a sizable reduction in this "age premium" from 1997 to 2007, and a slight increase at the end of the period when, as shown on the top right graph, the employment share of young workers decreased.

Sectors: the special role of construction. We next document sector-specific employment and earnings. The upper panel in Figure 7 shows the evolution of employment shares by sector. To facilitate interpretation we have aggregated sectors into 4 broad categories: industry (other than construction), construction, private services, and public services.[22] The graphs show two salient facts. The first one is the decline of industry

[22]Note that public employees in our data set belong to the general regime of the social security administration. Hence, some government employees, such as the armed forces or the judicial power, are not included.

Figure 7: Employment shares and earnings ranks, by sector



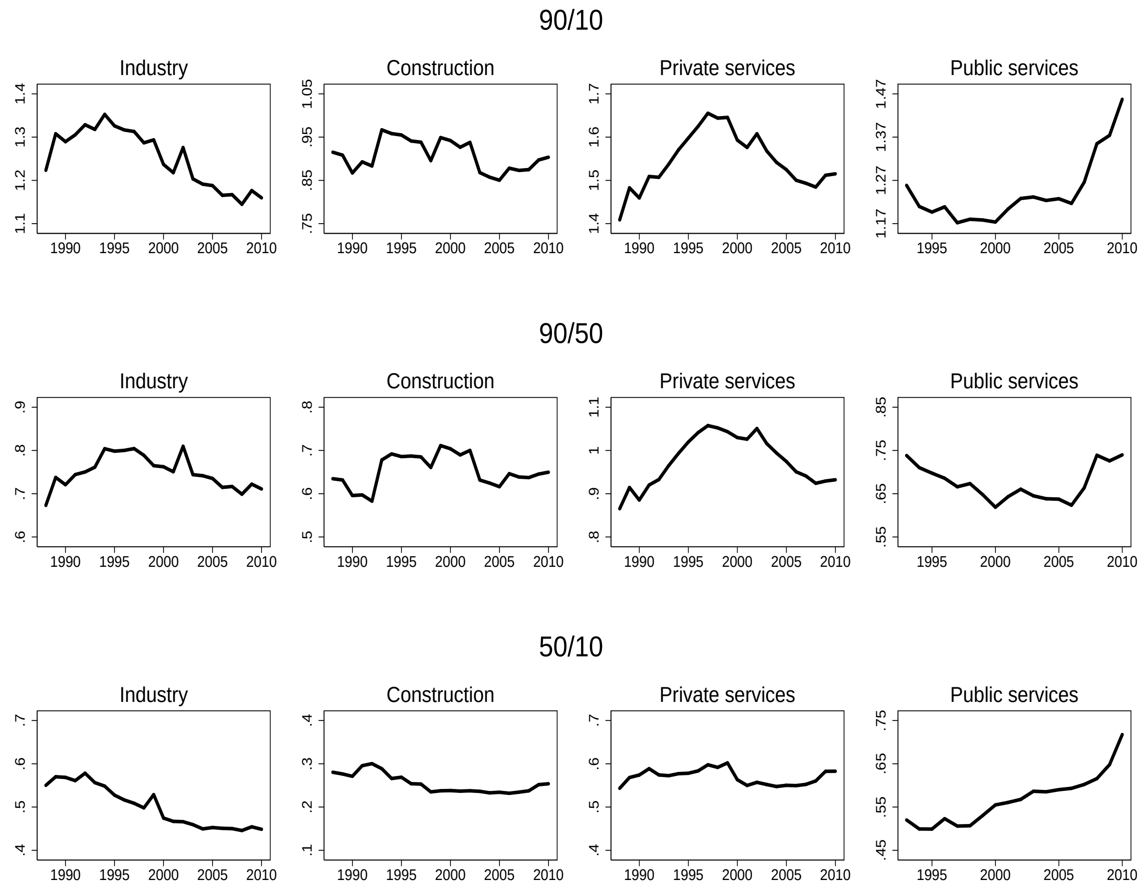
Notes: Source Social Security data. The upper graphs show employment shares, by sector. The lower graphs show sector-specific averages of ranks of daily earnings in the aggregate distribution. See Table A3 in the appendix for a definition of sectors.

in Spanish employment.[23] The second fact is the procyclical evolution of the share of construction. Between 1997 and 2007, the share of construction in male employment increased from 14% to 21%. That share then sharply decreased to 13% in 2010, less than its 1990 level. This remarkable evolution points to a special role of the construction sector in the Spanish economy. By comparison, private services experienced a steady increase during the whole period, and the employment share of public services decreased during the expansion before increasing in the recent recession.

The lower panel in Figure 7 shows that earnings in the construction sector increased during the period, particularly during the expansion episode. In 1988, the rank of a construction worker in the aggregate earnings distribution was 33% on average, while in 2010 the average rank was 42%. Half of the increase occurred between 1997 and 2006. By comparison, earnings ranks in industry and private services remained stable during the

[23]Despite the relative fall in manufacturing employment, in absolute numbers employment in that sector increased between 1995 and 2007. This contrasts with the continued decline in manufacturing employment in the US throughout the period.

Figure 8: Log-percentile ratios, by sector

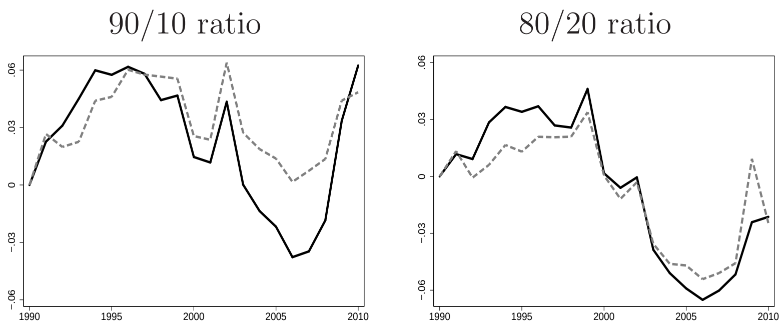


Notes: Source Social Security data. Log-ratios of estimated unconditional quantiles of daily earnings.

whole period, while public services experienced a marked increase in earnings during the expansion (a period where the employment share of public services decreased, as shown by the top right graph). Comparing the earnings ranks of construction workers with the sector shares suggests that demand for construction workers was high during the boom. The evidence is also suggestive of a negative demand shock during the bust, although relative earnings in construction did not fall after 2008.[24] Note also that the demand for construction workers during the expansion went in parallel with the fall in the college premium documented in Figure 6. This evolution contrasts with that of other Western countries such as the US, where high-skilled workers have been in high demand for the

[24]Downward wage rigidity might partly explain why relative earnings have not adjusted in the recession.

Figure 9: Log-percentile ratios, with and without the construction sector



Notes: Source Social Security data. Solid lines are ratios of estimated unconditional quantiles of daily earnings, dashed lines are ratios of estimated unconditional quantiles of daily earnings in a sample without the construction sector (both in logs, index zero at the start of the period). The tobit model for censoring correction is separately estimated in the sample without construction workers.

last three decades.[25]

Figure 8 shows the evolution of the 90/10, 90/50, and 50/10 log-percentile ratios within each of the four sectors considered.[26] The results indicate that, during the expansion, inequality decreased within both industry and private services. In contrast, in the recent recession within-sector inequality remained flat in all sectors but public services where the increase was substantial. The decomposition exercises that we report next aim at assessing the impact of all these different dimensions (employment composition, and between and within age/skill/sector earnings dispersion) on the evolution of inequality.

Before turning to the decomposition exercises, we first informally check the influence of the construction sector on the evolution of male inequality by reporting inequality measures in a sample without construction in Figure 9. The latter (indicated by dashed lines) exhibit a countercyclical evolution over the period, suggesting that the fluctuations of the construction sector alone cannot explain the Spanish pattern of inequality.[27] At

[25] To provide a finer view of sectoral differences, in the supplementary appendix we report percentage changes in sector-specific employment for a list of 50 disaggregated sectors (Table S5). The sectors are ranked by employment changes between 1997 and 2006 (left column) and between 2007 and 2010 (right column). The cyclicality of construction-related sectors is apparent from the table. During the expansion, among the 10 sectors with the largest percentage gains in employment, 4 sectors were construction-related. Other sectors whose employment shares increased substantially during the period were computer services, R&D, and advertising, for example. During the recent recession, in contrast, out of the 10 sectors whose percentage losses in employment have been the largest, 8 are directly or indirectly (e.g., cement or brick manufacturing) linked to construction.

[26] In the supplementary appendix we show the 80/20, 80/50, and 50/20 log-percentile ratios (Figure S14).

[27] As a check, we replicated the same exercise, while also taking out construction-related sectors such

the same time, the figure shows that the fall in inequality during the Spanish expansion, and the increase during the recent recession, are less pronounced in the sample without construction.[28]

5.2 Decomposition exercises

Methodology. The methodology we use is closely related to the decomposition approach in Autor *et al.* (2005).[29] We decompose the change in inequality between two periods, say t and t' ($t' > t$), into three components: change in composition, change in between-group prices, and change in within-group prices. To fix ideas, we present the approach in the case where skills and experience are the characteristics of interest.

We start by setting the notation. Let $\widehat{p}_{c,t}$ denote the size of a skill/experience cell c at time t, where we have indicated the time subscript for clarity. Let $\widehat{\mu}_{c,t}$ denote an estimate of the mean of log-earnings in cell c at time t. Let also $\widehat{F}_{c,t}$ be an empirical counterpart of the conditional cdf of de-meaned log-earnings $\log w_{it} - \widehat{\mu}_{c,t}$, for individual i in cell c. Finally, let $\widehat{p}_t$, $\widehat{\mu}_t$, and $\widehat{F}_t$ include all $\widehat{p}_{c,t}$, $\widehat{\mu}_{c,t}$, and $\widehat{F}_{c,t}$ for all skill/experience cells, respectively.

With this notation at hand, the earnings percentile ratio $R_{\tau,t}$ at time t can be written as:

$$R_{\tau,t} = R_\tau\left(\widehat{p}_t, \widehat{\mu}_t, \widehat{F}_t\right).$$

Similarly as Autor *et al.* (2005), we decompose the log-difference in inequality between t

as manufacture of bricks or cement and rental activities. The evidence in the supplementary appendix (Figure S15) shows a picture comparable to Figure 9.

[28]The 90/10 ratio decreases by 5.6 log points between 1997 and 2006, as opposed to 9.6 when including construction, while it increases by 4.1 log points in the sample without construction between 2007 and 2010, as opposed to 9.7 in the original sample. The 80/20 ratio is more similar between the two samples: it decreases by 7.5 log points when removing construction as opposed to 9.2 log points in the full sample, and then increases by 2.6 log points instead of 3.9. In addition, in the supplementary appendix (Figure S16) we reproduce Figure 1 in a sample without construction workers. The results show some evidence of a non-monotonic pattern, especially when comparing the 1993 recession to the expansion. However, the relationship is less clearly apparent than in Figure 1. Lastly, we also report log-percentile ratios in a sample without public services (Figure S17). We find that the inclusion of public services tends to slightly *reduce* the cyclical fluctuations of earnings inequality.

[29]As noticed by Autor *et al.*, this approach is also closely related to other decomposition methods proposed in the literature; see Juhn *et al.* (1993), DiNardo *et al.* (1996), and Lemieux (2008a), for example.

and t' as follows:

$$\begin{aligned}\log R_{\tau,t'} - \log R_{\tau,t} &= \underbrace{\log R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_t, \widehat{F}_t\right) - \log R_\tau\left(\widehat{p}_t, \widehat{\mu}_t, \widehat{F}_t\right)}_{\text{“composition effect”}} \\ &\quad + \underbrace{\log R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_{t'}, \widehat{F}_t\right) - \log R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_t, \widehat{F}_t\right)}_{\text{“between-group price effect”}} \\ &\quad + \underbrace{\log R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_{t'}, \widehat{F}_{t'}\right) - \log R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_{t'}, \widehat{F}_t\right)}_{\text{“within-group price effect”}}.\end{aligned}$$

The inequality measures $R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_t, \widehat{F}_t\right)$ and $R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_{t'}, \widehat{F}_t\right)$ correspond to two counterfactual earnings distributions. The first one is simply obtained by re-weighting the time-t conditional quantiles

$$w_{c,t}^q = \exp\left(\widehat{\mu}_{c,t} + \widehat{F}_{c,t}^{-1}(q)\right)$$

by the proportions of skill/experience cells at time t'. The composition effect thus represents the change in inequality due to changes in the composition of employment only.[30] The second counterfactual distribution is obtained by re-weighting the following counterfactual conditional quantiles

$$w_{c,t}^{q,BG} = \exp\left(\widehat{\mu}_{c,t'} + \widehat{F}_{c,t}^{-1}(q)\right),$$

using skill/experience cells at time t'. The between-group price effect thus represents the change in inequality due to changes in cell-specific means of log-earnings. The within-group price effect then captures the impact of changes in cell-specific distributions, keeping cell means constant.

In practice, the counterfactual inequality measure $R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_t, \widehat{F}_t\right)$ is calculated by re-weighting cell-specific log-earnings at time t by the cell sizes at time t'. In case log-earnings are top or bottom-coded, we use predicted values from the tobit model. The second counterfactual inequality measure $R_\tau\left(\widehat{p}_{t'}, \widehat{\mu}_{t'}, \widehat{F}_t\right)$ is obtained by shifting time-t log-earnings in each cell c by the mean difference $\widehat{\mu}_{c,t'} - \widehat{\mu}_{c,t}$, and re-weighting the cell-specific observations by cell sizes at time t'.

This decomposition can be performed using different characteristics to form the cells. We will use skill/experience cells, as well as skill/experience/sector cells based on our broad 4-sector classification. Note also that the results depend on the order of the decomposition: composition effect, between-group price effect, and within-group price effect, in

[30]Note that this type of decomposition relies on a partial equilibrium assumption according to which quantities of skill/experience do not affect prices.

Table 2: Decomposition of inequality changes: employment composition and price effects

	Age and occupation groups				Age, occupation, and sectors			
	1988-1996							
	Total	Composition	Between	Within	Total	Composition	Between	Within
90/10	10.81	3.30	6.37	1.14	10.92	2.80	7.56	0.56
90/50	11.03	1.95	4.73	4.35	11.23	2.17	4.09	4.97
50/10	-0.22	1.36	1.64	-3.21	-0.30	0.64	3.47	-4.41
	1997-2006							
	Total	Composition	Between	Within	Total	Composition	Between	Within
90/10	-9.58	-3.20	-1.80	-4.57	-10.61	-3.47	-3.31	-3.83
90/50	-3.61	-2.11	-0.57	-0.93	-4.86	-0.18	-3.57	-1.12
50/10	-5.96	-1.09	-1.23	-3.64	-5.76	-3.30	0.26	-2.71
	2007-2010							
	Total	Composition	Between	Within	Total	Composition	Between	Within
90/10	9.71	5.65	-0.31	4.37	7.96	7.90	0.08	-0.02
90/50	1.55	3.17	-0.50	-1.11	0.89	2.76	-0.79	-1.08
50/10	8.15	2.48	0.19	5.48	7.07	5.14	0.87	1.06

Notes: Differences in log-percentile ratios. Decomposition of the total change into composition effect, between-group price effect, and within-group price effect. On the right panel, the sample is smaller as a result of missing sector data for some observations (3.2%).

this order. We checked that the results remain qualitatively similar when changing the order of the decomposition.

Decomposition results: 1988-1996. Table 2 shows the results of two decompositions: using age/occupation cells only (columns 1-4), and using age/occupation/sector cells (columns 5-8).[31] Let us first consider the 1988-1996 period. Over this period, the 90/10 earnings percentile ratio increased by 10.8 log points. Out of this, 3.3 log points are due to composition changes. This means that, if employment composition in terms of occupation and age groups had been constant to its 1996 level, inequality would have increased by 7.5 log points only. Changes in between-group prices (that is, changes in cell-specific means of log-earnings) explain a 6.4 log points increase. The remaining 1.1 log point is due to changes in cell-specific distributions keeping the means constant (that is, to changes in within-group prices). Thus, 60% of the total increase in inequality during the 1988-1996 period is due to between-group price effects, and 30% is due to composition effects.

Between-group price effects in part reflect the large increase in the skill premium around the 1993 recession, when skill and age composition did not vary much (see Figure

[31] When including sector dummies as covariates, we lose 3.2% of observations due to missing data. As a result, total inequality changes differ somewhat between columns 1 and 5 in Table 2.

6). Moreover, the results show that between-group price effects mostly affected upper-tail inequality. In contrast, while within-group price effects contributed to the increase in the 90/50 ratio, they pushed towards a reduction in 50/10 inequality. This is because, in this period, changes in within-cell earnings distributions had a positive effect on the two tails of the overall distribution, but had a negative effect on the median. In addition, when comparing with the right panel of the table, we see that allowing for the sectoral dimension has no major effect on the decomposition in the 1988-1996 period. One difference is that between-group price effects explain a larger share of the changes in lower-tail inequality when including sectors. The small role of within-group price effects (0.6 log points) is in line with the absence of a clear pattern in within-sector 90/10 inequality in the first subperiod (see Figure 8).

Decomposition results: 1997-2006. Turning to the 1997-2006 period, Table 2 shows that the 9.6 log-points decrease in 90/10 inequality can be decomposed into 3.2 log-points due to composition changes, 1.8 due to between-group price effects, and 4.6 due to within-group price effects. Composition changes mostly affect upper-tail inequality. The between-group price effects may be due to the decrease in the earnings gap between older and younger workers during the expansion, a period when the skill premium did not show a clear trend (see Figure 6). However, between-group price effects only account for 20% of the inequality change, compared to 50% for within-group price effects.

When accounting for sectors in addition to age and occupation, composition changes then explain a substantial part (3.3 log points, that is 60%) of the decrease in lower-tail inequality between 1997 and 2006. This reflects changes in sector shares during the period, in particular the relative decrease of industry and the growth of construction (see Figure 7). Moreover, between-group price effects then explain most of the fall in 90/50 inequality (3.6 log points), reflecting the evolution of average earnings in age/occupation/sector cells. At the same time, note that within-group price effects– which is the part left unexplained when accounting for composition changes and changes in average cell-specific earnings– remain substantial even when accounting for age, occupation and sectors. This result is consistent with Figure 8, which shows that within-sector inequality decreased in the expansion, particularly in industry and private services. The magnitude of within-group price changes suggests that other factors, not accounted for in this analysis, may have contributed to the fall in inequality during the expansion episode.

Decomposition results: 2007-2010. The bottom panel in Table 2 shows the results for the 2007-2010 period. When accounting for age and occupation only, within-group price effects explain more than 40% of the total inequality increase. Interestingly, when adding sectors to the decomposition, composition effects fully explain the evolution of the 90/10 ratio. Looking at upper and lower-tail inequality shows that the main difference concerns the 50/10 ratio. This suggests that most of the within-group price changes estimated on the left graph for the 2007-2010 period reflect changes in sectoral composition, within age/occupation cells. This is consistent with Figure 7, which shows strong changes in sectoral shares during the recent recession, particularly for construction. In this period, price effects seem at best modest, possibly reflecting the fact that wages take time to adjust. The results thus show an interesting contrast between the 1993 recession, when price effects seem to have played a major role, and the recent recession, when composition effects (and particularly changes in sectoral composition) have been dominant.[32]

Additional decomposition exercises. In order to better understand the nature of changes in employment composition, we performed a related exercise keeping the relative employment shares of all sectors but construction constant.[33] Composition effects calculated in this way reflect changes in the employment shares of age and occupation groups, as well as in the share of the construction sector. These "construction-driven" composition effects explain little of the evolution of upper-tail inequality. In contrast, they appear to have been a major driver in the recent evolution of lower-tail inequality. Indeed, they explain more than half of the decrease in the 50/10 ratio in the 1997-2006 expansion, close to the share explained by all sectors and age/occupation groups. In the 2007-2010 recession, construction-driven composition effects explain 50% of the increase in lower-tail inequality, compared to 75% when accounting for all sectors.

Lastly, we replicated the first decomposition exercise using education instead of occupation as a proxy for skills.[34] The results of the decomposition are similar to Table 2 for the 2007-2010 period. However, the other two periods show some differences. In partic-

[32] In the supplementary appendix we show the results of the same two decompositions, using the 80/20, 80/50, and 50/20 ratios as inequality measures (see Table S6). The results show similar patterns as Table 2, changes being of smaller magnitudes. One difference concerns upper-tail inequality in the expansion as, when accounting for age skills and sectors, between-group price effects are comparatively much smaller for 80/50 than for 90/50 inequality (1.2 log points versus 3.6). Another difference concerns the 2007-2010 period: unlike the 90/10 ratio, 80/20 inequality is almost fully explained by composition changes when accounting for age and occupation groups.

[33] See the supplementary appendix (Figure S18).

[34] See the supplementary appendix (Table S7).

ular, when relying on education instead of education, the fall in inequality between 1997 and 2006 is in an important part attributable to changes in between-group prices. This may reflect the decline in the college premium documented in Figure 6. Nevertheless, a note of caution is in order when interpreting these results, due to the lower quality of the education data.

5.3 Other factors

In the last part of the section we extend the analysis to account for three potentially important factors in the Spanish case: the duality of labor contracts, immigration, and the minimum wage.

Labor market duality. After their introduction in 1984, the use of temporary contracts grew rapidly up to approximately 33% of the labor force by the early 1990s. The proportion has remained relatively stable since then until the current crisis, and it represents the largest share in Europe. Most of the literature has focused on the determinants of the duration and conversion rates of temporary contracts into permanent positions,[35] or on the effect of dual employment protection on productivity (Dolado *et al.*, 2011). However, less is known about the effect of temporary contracts on earnings over time.

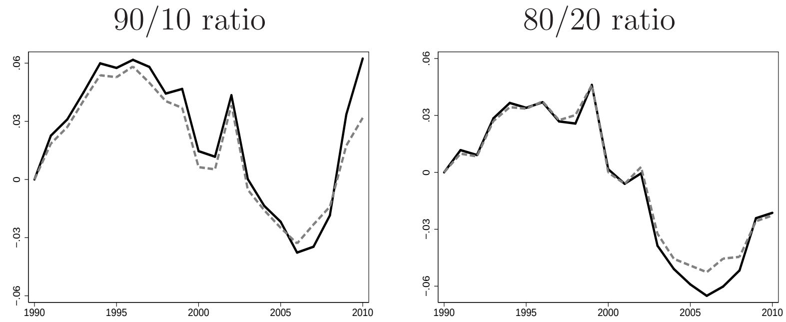
In our administrative data, reliable information regarding the type of contract (permanent versus temporary) is available only since 1998, thus we restrict this analysis to the 1998-2010 subperiod. The left panel of Figure A1 in the appendix reports the evolution of the employment share of temporary workers in our data. Temporary contracts are highly concentrated among the young, immigrants, and low-skilled workers. By sector, the proportion of temporary contracts is disproportionately high in construction: 63% on average over the period, and 67% from 1998 to 2006.[36] The right panel of Figure A1 shows the evolution of median earnings of permanent and temporary workers. The ratio between permanent and temporary median earnings fell by almost 20% between 1998 and 2007. It then increased by about 7% from 2007 to 2010.

The evolution of between-type-of-contract inequality is consistent with the evolution of total earnings inequality between 1998 and 2010. In addition, given the high share of temporary contracts in the construction sector, this evolution may partly reflect the surge and subsequent fall in demand for construction workers. To get additional insight, we also

[35] See for example Amuedo-Dorantes (2000), Güell and Petrongolo (2007), and García-Pérez and Muñoz-Bullón (2010).

[36] See the supplementary appendix (Figure S19).

Figure 10: Log-percentile ratios, with and without immigrant workers



Notes: Source Social Security data. Solid lines are ratios of estimated unconditional quantiles of daily earnings, dashed lines are ratios of estimated unconditional quantiles of daily earnings in a sample without immigrants (both in logs, index zero in 1990). The tobit model for censoring correction is separately estimated in the sample without immigrants.

performed a decomposition exercise similar to above, using age, occupation, sectors, and a binary indicator of type of contract to form the cells. We found that including the type of contract had a very small impact on the results.[37]

Immigration. During the last decade the inflows of immigrants in Spain increased sharply. Due to illegal immigration, available data sources (population census, administrative registers of residence and work permits, labor force survey...) do not always coincide in the measurement of the stock of foreign population in Spain. Similarly, our data set only contains immigrants registered with the social security administration. As shown on the left panel of Figure A2 in the appendix, the proportion of foreign-born workers among male employees increased from 5% in 2000 to 16.4% in 2007, and then decreased to 14.4% in 2010. So, according to our data the period of fall in inequality was associated with increased immigration, while the recent period of inequality increase has been associated with decreasing immigration. In addition, the right panel of Figure A2 shows that, during the same period the native-immigrant earnings gap experienced only minor changes until 2007, while it increased in the recent recession.

As a crude way of assessing the effect of immigration of inequality, Figure 10 shows the evolution of the earnings log-percentile ratios in a sample without immigrants. We see that inequality levels are similar to the ones in the full sample until 2005, suggesting that immigration had a small effect on aggregate earnings inequality. From 2006 onward,

[37]See the supplementary appendix (Table S8). When interpreting these results, the partial equilibrium nature of the decomposition exercise is worth keeping in mind.

in contrast, removing immigrants from the sample tends to lower the cyclical fluctuations of inequality, suggesting that the presence of immigrants contributed to accentuate the recent inequality increase. Indeed, when removing immigrant workers the 90/10 ratio increases by 5.5 log points between 2007 and 2010, compared to a 9.7 log points increase in the full sample.

One limitation of the exercise is that immigration could have had an effect on earnings of non-immigrants, for example by reducing the wages of natives working in similar occupations.[38] In addition, the patterns in Figure 10 could partly reflect the fact that male immigrants are highly concentrated in construction.[39] To assess whether including immigration as another factor has an impact on composition and price effects, we performed an additional decomposition exercise, using age, occupation, sectors, and a native/immigrant binary indicator. The results show small changes relative to Table 2, especially when focusing on the 2007-2010 period.[40]

Minimum wage. In the US, several studies have argued that the decline in the Federal minimum wage partly explains the increase in earnings inequality in the 1980s (see e.g. DiNardo *et al.*, 1996, Lee, 1999). The minimum wage is unlikely to have played a major role in the evolution of Spanish earnings inequality, however. Most of the 1998-2006 period was characterized by a slight decrease in the real value of the minimum wage, while the end of the 2000s saw a marked increase, by 18% between 2004 and 2009 in real terms.[41] This timing is unable to explain the patterns of total and lower-tail inequality that we document. If anything, the minimum wage increase at the end of the period may have pushed inequality in the opposite direction.

Discussion. A close look at other factors confirms that the Spanish expansion was a period of high demand for certain types of workers: the low-skilled, temporary workers, and immigrants, all well-represented in the construction sector. The decomposition exercises suggest that demand-driven employment and wage increases quantitatively contributed to the fall in inequality. The expansion episode contrasts with the earlier period where price effects, reflecting demand for high-skilled workers, had a large impact. The recent

[38] Carrasco *et al.* (2008) do not find significant effects of immigration on either the employment rates or the wages of native workers during the second half of the 1990s.

[39] See the supplementary appendix for shares of foreign-born workers in employment, by sector (Figure S20).

[40] See the supplementary appendix (Table S9).

[41] See the supplementary appendix (Figure S21).

recession shows an opposite evolution compared to the previous decade, with a sharp drop in demand for the workers who most benefited from the boom. Overall, these findings highlight the sensitivity of the Spanish economy to business conditions, and the fact that lower-middle wage workers have been most affected by them.

6 Unemployment-adjusted inequality

The previous sections have provided evidence on earnings inequality while abstracting for unemployment. For example, monthly and daily earnings were computed for individuals having worked at some point during a particular month. In this last section, we aim at documenting the evolution of inequality in the broader population, including employed and unemployed individuals. For this purpose, as a complement to the main analysis, we document the evolution of unemployment-adjusted inequality measures obtained by imputing income values to the unemployed.

6.1 Imputation methods

We compare and contrast two different approaches to impute earnings values to the unemployed.

Approach 1: Potential earnings. Our first approach is based on a neoclassical Mincer model where potential earnings are equal to the marginal productivity of labor. As in Heckman (1979), individuals decide whether or not to work by comparing their potential earnings with their reservation wage. Several methods have been proposed to account for non-random selection into employment in this framework.[42] We follow Olivetti and Petrongolo (2008) and make use of the panel dimension of our data. For each unemployed worker, we recover his daily earnings observation from the nearest wave where he is working.[43] Hence, when unemployment spells are preceded and followed by two employment relationships, the imputed earnings follow a step function with a jump in the middle of the spell. The underlying assumption is that the latent earnings of an individual can be proxied by his earnings in the nearest wave where he is employed. Note that this method is based on longitudinal earnings information, and thus effectively allows for selection on unobservables.

[42] See Neal (2004) and Blundell *et al.* (2007) for recent examples.

[43] When the earnings observation is censored, we use predicted values from the tobit model. We proceed similarly in approach 2 below.

Approach 2: Unemployment benefits. One limitation of the previous approach is that it is not directly related to the benefits individuals actually receive when unemployed. As a complement, our second approach uses unemployment benefits to impute labor income to the unemployed. We use a simple approximation that mimics the benefits rules of the Spanish system over the period.[44] In this second approach we also use the panel structure of the data to compute the duration of the unemployment spell. Our measure of previous earnings is the last daily earnings that the individual received when he was working. One specific feature of this approach is that benefits decrease with the duration of unemployment.[45]

6.2 Results

We start by describing the evolution of inequality in potential earnings, as shown on the top panel of Figure 11.[46] Here we focus on daily measures for comparison with the main results. Annual earnings inequality measures based on imputed values show a qualitatively similar evolution, as we document in the supplementary appendix (Figures S23 and S24). The top panel of Figure 11 shows that the level of inequality in potential earnings is higher than that of earnings inequality conditional on employment. This reflects the fact that selection in employment is positively correlated with potential earnings. However, the overall pattern of evolution is preserved.

We next turn to our second approach to impute income values to the unemployed, based on the benefits rule. By construction, this approach takes into account the duration of unemployment. The bottom panel of Figure 11 shows that the level of inequality is substantially higher than when using the potential earnings method. In terms of evolution, the 2008-2010 recession seems to have had a smaller effect relative to the recession of the early 1990s. This could be due to the fact that these numbers partly reflect the duration of unemployment spells, so the effect of a recession on inequality may take some time to appear. In contrast, the fall in inequality during the expansion period is substantial.

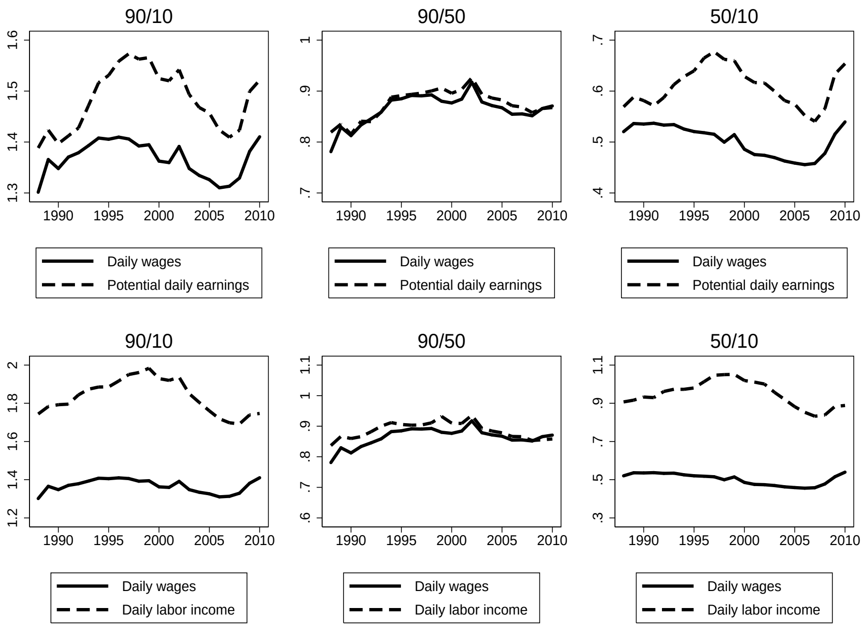
These results show that the qualitative patterns of male inequality are preserved when accounting for unemployment. However, the quantitative conclusions on the level and

[44] As a simplification, we assume that all workers are eligible to unemployment benefits, and that the benefits rule is stationary. The percentage of previous earnings imputed depends on the number of months in unemployment as follows: 70% (1-6 months), 60% (6-24), 50% (25-48), 40% (49-72), 30% (73-96), 20% (97-120), and 10% (>120).

[45] Spain presents high cyclical variations in employment and high incidence of long-term unemployment, see the supplementary appendix (Figure S22).

[46] In the supplementary appendix we report the estimated quantiles for selected years (Table S10).

Figure 11: Unemployment-adjusted inequality measures: Log-percentile ratios of daily wages, potential earnings, and labor income



Notes: Source Social Security data. Solid lines are ratios of estimated unconditional quantiles of daily earnings conditional on employment (in logs). Dashed lines are ratios of estimated unconditional quantiles of potential daily earnings (top panel), and ratios of estimated unconditional quantiles of daily labor income based on imputed unemployment benefits (bottom panel).

evolution of inequality appear sensitive to the imputation method. In high-unemployment countries such as Spain, it seems worth further exploring combined inequality measures that take the impact of unemployment into account, in order to better assess the welfare consequences of inequality.

7 Conclusion

In this paper we use administrative data from the social security to characterize the evolution of earnings inequality and employment in Spain from 1988 to 2010. We document that the dispersion of the earnings distribution experienced substantial changes over the past two decades. Male inequality fell during the expansion, and increased sharply in the two recessions. The magnitudes of these changes over the cycle of earnings inequality are large by international standards. This evolution is partly explained by sizable changes in the composition of employment in terms of occupation, age groups, and sectoral composition. In particular, we find that the inequality increase during the recent recession is fully accounted for by composition effects. During the whole period, the evolution of earnings inequality went in parallel with the fact that cyclical employment fluctuations mostly affected the lower-middle part of the distribution.

The construction sector appears to have played a special role in this evolution. The Spanish boom of the late 1990s and 2000s was also a housing boom. Driven by fluctuations in demand, employment of construction workers rose, and subsequently fell during the housing bust. In turn, these movements contributed to the countercyclical evolution of inequality. This suggests that policies that fostered the demand for housing had sizable effects on labor market outcomes. More generally, our findings motivate further studies of the interactions between the housing market and the labor market, in the US but also in other countries that have experienced strong housing booms and busts such as the UK, Ireland, or Denmark.

Documenting female earnings inequality in Spain is another important avenue for future work. In the social security data we found that, for females, 90/10 inequality increased by more than 15 log-points between the early 1990s and the early 2000s. The end of the period shows a countercyclical pattern, albeit less pronounced than for males. Even though these patterns are suggestive, the design of the data set prevents us from drawing firm conclusions. In Bonhomme and Hospido (2013) we use the tax data, which have a proper panel structure, to compare male and female earnings distributions in the more recent period. Providing a longer view on female inequality might require different data.[47]

[47] Another important limitation of the social security sample is that it is silent on the evolution of the right tail of the earnings distribution. Alvaredo and Saez (2009) use tax data to document the evolution of top income shares in Spain over the last century. See also Bonhomme and Hospido (2013) for the more recent period.

References

[1] Abadie, A. (1997), "Changes in Spanish labor income structure during the 1980's: a quantile regression approach", *Investigaciones Económicas*, XXI(2), 253-272.

[2] Acemoglu, D. (2002), "Technical Change, Inequality and the Labor Market", *Journal of Economic Literature*, 40, 7-72.

[3] Alvaredo, F., and E. Saez (2009), "Income and Wealth Concentration in Spain from a Historical and Fiscal Perspective", *Journal of the European Economic Association*, 7, 1140-1167.

[4] Amuedo-Dorantes, C. (2000), "Work transitions into and out of involuntary temporary employment in a segmented market: evidence from Spain", *Industrial and Labor Relations Review*, 53(2), 309-325.

[5] Aparicio, A. (2010), "High-School Dropouts and Transitory Labor Market Shocks: The Case of the Spanish Housing Boom", IZA discussion paper 5139.

[6] Arranz, J. M., and C. García-Serrano (2013), "Earnings differentials and the changing distribution of wages in Spain, 2005-2010", *International Journal of Manpower*, forthcoming.

[7] Autor, D. H., L. F. Katz, and M. S. Kearney (2005), "Residual Wage Inequality: The Role of Composition and Prices", NBER working paper 11628.

[8] Autor, D. H., L. F. Katz, and M. S. Kearney (2008), "Trends in U.S. Wage Inequality: Re-assessing the Revisionists", *Review of Economics and Statistics*, 90, 300-323.

[9] Autor, D. H., F. Levy, and R.J. Murnane (2003), "The Skill Content of Recent Technological Change: An Empirical Exploration", *Quarterly Journal of Economics*, 118, 1279-1334.

[10] Ayuso, J. and F. Restoy (2007), "House prices and rents in Spain: does the discount factor matter?", *Journal of Housing*, 16, 291-308.

[11] Barlevy, G., and D. Tsiddon (2006): "Earnings Inequality and the Business Cycle," *European Economic Review*, 50(1), 55-89.

[12] Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007), "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds", *Econometrica*, 75, 323-363.

[13] Bonhomme, S., and L. Hospido (2013), "Earnings inequality in Spain: new evidence using tax data", *Applied Economics*, 45, 4212-4225.

[14] Boudarbat, B., T. Lemieux, and W. C. Riddell (2006), "Recent Trends in Wage Inequality and the Wage Structure in Canada", in D. A. Green, and J. R. Kesselman (Eds.): Dimensions of Inequality in Canada, Vancouver, BC: UBC Press.

[15] Bound, J. and G. Johnson (1992), "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations", *American Economic Review*, 82, 371-392.

[16] Bover, O., S. Bentolila, and M. Arellano (2002), "The Distribution of Earnings in Spain During the 1980s: The Effects of Skill, Unemployment, and Union Power" in D. Cohen, T. Piketty and G. Saint-Paul (Eds.): The Economics of Rising Inequalities, Oxford University Press and CEPR, Chapter 1, 3-53.

[17] Card, D., J. Heining, and P. Kline (2013), “Workplace Heterogeneity and the Rise of West German Wage Inequality”, *Quarterly Journal of Economics*, 128, 967-1015.

[18] Carrasco, R., J. F. Jimeno, and A. Carolina Ortega (2008), “The effect of immigration on the labor market performance of native-born workers: some evidence for Spain”, *Journal of Population Economics*, 21, 627-648.

[19] Carrasco, R., J. F. Jimeno, and A. Carolina Ortega (2011), “Accounting for changes in the Spanish Wage Distribution: The Role of Employment Composition Effects”, Banco de España working paper 1120.

[20] Casado, J., and H. Simón (2013), “La evolución de la estructura salarial en España (2002-2010)”, mimeo.

[21] Charles, K., E. Hurst, and M. Notowidigdo (2013), “Manufacturing Decline, Housing Booms, and Non-Employment,” NBER Working Paper 18949.

[22] Del Río, C. and J. Ruiz-Castillo (2001), “Accounting for the decline in Spanish household expenditures inequality during the 1980s”, *Spanish Economic Review*, 3, 151-175.

[23] DiNardo, J. E., N. Fortin, and T. Lemieux (1996), “Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach”, *Econometrica*, 64, 1001-1044.

[24] Dolado J. J., C. Garcia-Serrano and J. F. Jimeno (2002), “Drawing lessons from the boom of temporary jobs in Spain”, *Economic Journal*, 112, F270-295.

[25] Dolado, J. J., S. Ortigueira, and R. Stucchi (2011), “Does dual employment protection affect TFP? Evidence from Spanish manufacturing firms”, mimeo.

[26] Dustmann, C., J. Ludsteck, and U. Schonberg (2009), “Revisiting the German Wage Structure”, *Quarterly Journal of Economics*, 124, 843-881.

[27] Farré, L. and F. Vella (2008), “Macroeconomic Conditions and the Distribution of Income in Spain”, *LABOUR: Review of Labour Economics and Industrial Relations*, 22, 383-410.

[28] Felgueroso, F., M. Hidalgo, and S. Jiménez-Martín (2010), “Explaining the fall of the skill wage premium in Spain”, FEDEA Annual Monograph Conference Talent, effort and social mobility.

[29] García-Montalvo, J. (2007), “Algunas consideraciones sobre el problema de la vivienda en España ”, *Papeles de Economía Española*, 113, 138-153.

[30] García-Pérez, J. I. and F. Muñoz-Bullón (2010), “Transitions into permanent employment in Spain: An empirical analysis for young workers”, *The British Journal of Industrial Relations*, forthcoming.

[31] Garriga, C. (2010), “The Role of Construction in the Housing Boom and Bust in Spain”, FEDEA Monograph on The Crisis of the Spanish Economy.

[32] Goldin, C. and L. F. Katz (1998), “The Origins of Technology-Skill Complementarity”, *Quarterly Journal of Economics*, 113, 693-732.

[33] González, L. and F. Ortega (2009), “Transitions Immigration and Housing Booms: Evidence from Spain”, CReAM Discussion Paper 19/09.

[34] Gosling, A., S. Machin, and C. Meghir (2000), “The Changing Distribution of Male Wages, 1966-1992”, *Review of Economic Studies*, 67, 635-666.

[35] Güell, M. and B. Petrongolo (2007), 'How binding are legal limits? Transitions from temporary to permanent work in Spain', *Labour Economics*, 14, 153-183.

[36] Guvenen, F., S. Ozkan, and J. Song (2014), "The Nature of Countercyclical Income Risk", *Journal of Political Economy*, 122(3), 621-660.

[37] Heathcote, J., F. Perri, and G. Violante (2010), "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States 1967-2006", *Review of Economic Dynamics*, 13, 15-51.

[38] Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47, 153-161.

[39] Hidalgo, M. (2008), "Wage Inequality in Spain 1980-2000", Universidad Pablo de Olavide, Economics Department working paper 08.08.

[40] Izquierdo, M. and A. Lacuesta (2012), "The contribution of changes in employment composition and relative returns to the evolution of wage inequality: the case of Spain", *Journal of Population Economics*, 25, 511–543.

[41] Jaimovich, N., and H. Siu (2013), "The Trend is the Cycle: Job Polarization and Jobless Recoveries," unpublished working paper.

[42] Juhn, C., K.M. Murphy, and B. Pierce (1993), "Wage Inequality and the Rise in Returns to Skill", *Journal of Political Economy*, 101, 410-442.

[43] Katz, L.F. and K.M. Murphy (1992), "Changes in Relative Wages, 1963–1987: Supply and Demand Factors", *Quarterly Journal of Economics*, 107, 35-78.

[44] Lacuesta, A., S. Puente, and E. Villanueva (2012): "The Schooling Response to a Sustained Increase in Low-skill Wages: Evidence from Spain 1989-2009," Banco de España working paper 1208.

[45] Lee, D. S. (1999), "Wage Inequality in the United States during the 1980s: Rising Dispersion or Falling Minimum Wage?", *Quarterly Journal of Economics*, 114, 977-1023.

[46] Lemieux, T. (2008a), "What do we Really Know about Changes in Wage Inequality?", mimeo.

[47] Lemieux, T. (2008b): "The Changing Nature of Wage Inequality," *Journal of Population Economics*, 21(1), 21-48.

[48] Levy, F. and R.J. Murnane (1992), "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations", *Journal of Economic Literature*, 30, 1333-1381.

[49] Manacorda, M. (2004): "Can the Scala Mobile Explain the Fall and Rise of Earnings Inequality in Italy? A Semiparametric Analysis, 1977-1993," *Journal of Labor Economics*, vol. 22, 585-613.

[50] Neal, D. (2004), "The Measured Black-white Wage Gap Among Women is Too Small", *Journal of Political Economy,* 112, S1-S28.

[51] Olivetti, C. and B. Petrongolo (2008), "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps", *Journal of Labor Economics*, 26, 621-654.

[52] Pijoan-Mas, J., and V. Sánchez-Marcos (2010), "Spain is Different: Falling Trends of Inequality", *Review of Economic Dynamics*, 13, 154-178.

[53] Piketty, T. and Saez, E. (2006): "The Evolution of Top Incomes: A Historical and International Perspective," *American Economic Review*, Papers and Proceedings, vol. 96, 200-205.

[54] Simón, H. (2009), "La desigualdad salarial en España: Una perspectiva internacional y temporal", *Investigaciones Económicas*, 33, 439-471.

APPENDIX

Table A1: Sample composition and descriptive statistics (men)

	Whole sample				Working individuals			
Individuals	52,878				52,599			
Observations	7,375,381				5,185,955			
	1988	1997	2007	2010	1988	1997	2007	2010
Age	37.02	37.86	38.16	38.95	37.52	37.92	38.12	38.96
	(8.20)	(8.35)	(8.14)	(8.02)	(8.21)	(8.33)	(8.14)	(8.02)
Immigrants (%)	1.54	3.95	16.85	17.43	1.57	3.62	16.53	14.54
Engineers-College	7.34	7.37	7.33	7.70	7.70	7.79	8.04	9.53
Technicians	3.50	3.72	4.50	4.64	3.95	4.53	5.00	5.74
Adm. managers	5.82	5.72	5.25	5.07	6.58	6.81	5.67	5.85
Assistants	4.42	4.25	3.63	3.63	5.12	5.26	4.07	4.32
Adm. workers	18.77	19.53	18.97	20.00	20.06	21.86	19.59	22.18
Manual workers	60.16	59.41	60.31	58.95	56.58	53.74	57.63	52.38
Annual days worked=0	14.41	25.25	14.51	19.18	-	-	-	-
Top-coded	-	-	-	-	27.10	21.48	16.07	17.98
Bottom-coded	-	-	-	-	3.01	4.00	3.37	4.96
Median daily earnings	-	-	-	-	46.54	48.36	50.93	53.76
	-	-	-	-	(19.8)	(23.6)	(23.9)	(24.7)
Temporary (%)	-	-	-	-	-	32.17	28.57	24.86

Notes: Standard deviations of non-binary variables in parentheses.

Table A2: Estimated quantiles of daily earnings and percentile ratios

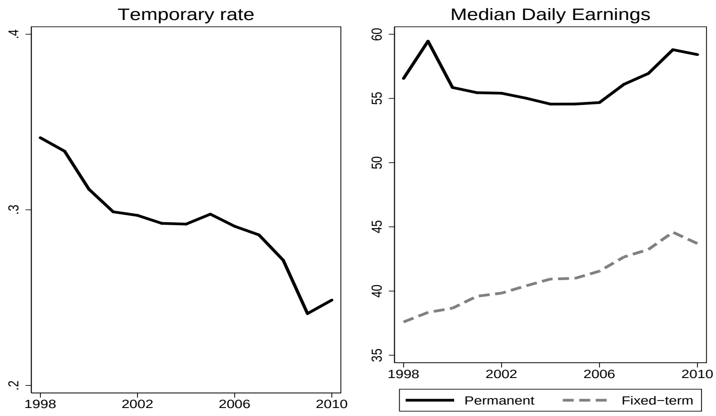
	1988	1997	2007	2010	1988-1996 Δlog (×100)	1997-2006 Δlog (×100)	2007-2010 Δlog (×100)
(A) Estimated quantiles of daily earnings							
w^{10}	27.7	28.9	32.2	31.3	3.91	8.26	-2.75
w^{50}	46.5	48.4	50.9	53.8	3.69	2.30	5.40
w^{90}	101.6	117.9	119.8	128.4	14.72	-1.32	6.96
(B) Percentile ratios							
w^{90}/w^{10}	3.67	4.08	3.72	4.10	10.81	-9.58	9.71
w^{90}/w^{50}	2.18	2.44	2.35	2.39	11.03	-3.61	1.55
w^{50}/w^{10}	1.68	1.67	1.58	1.72	-0.22	-5.96	8.15

Notes: Unconditional quantiles estimated from Social Security data.

Table A3: Sectors definitions

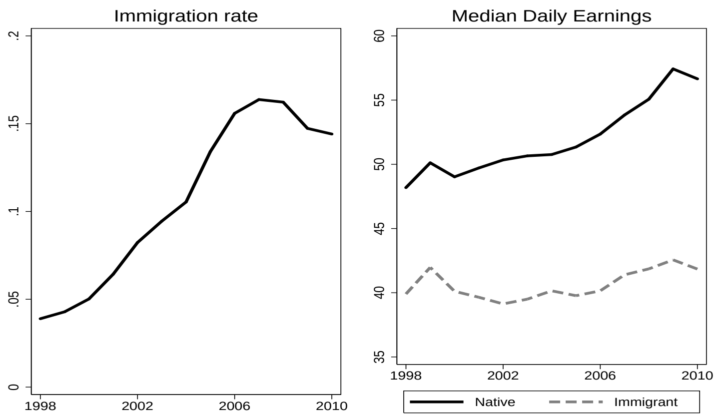
Industry:	Agriculture, forestry and fishing, mining and quarrying, Manufacture of food, beverages, tobacco, textiles, wood, paper, coke, chemicals, plastic and ceramic products, glass, cement, metals, machinery and equipment, electronic products, motor vehicles, furniture and other manufacturing.
Construction:	All general building works, installation systems and extensions (electrical system, painting, plumbing and tiling, carpentry, flooring, plastering), civil engineering works, renting of the building equipment.
Services:	Sales, accommodation, storing, transport, telecommunications and energy, financial services, corporate and personal services, public administration, education, health, social activities. **Public services:** When the employer is any local, regional or national government institution. **Private services:** Otherwise.

Figure A1: Temporary employment and median earnings



Notes: Source Social Security data. The left graph shows the share of temporary/fixed-term contracts in employment. The right graph shows median earnings by type of contract.

Figure A2: Immigration employment and median earnings



Notes: Source Social Security data. The "immigration rate" is computed as the share of foreign-born workers among employees.