Recent changes in British wage inequality: Evidence from firms and occupations

Schaefer, Daniel and Singleton, Carl

University of Edinburgh

January 2017

Recent changes in British wage inequality:
Evidence from firms and occupations[*]

Daniel Schäfer[†]   Carl Singleton[‡]

Latest version: January 2017

Abstract

Using a dataset covering a large sample of employees and their mostly very large employers, we study the dynamics of British wage inequality over the past two decades. Contrary to other studies, we find little evidence that recent increases in inequality have been driven by differences in the average wages paid by firms. Instead greater dispersion within firms can account for the majority of changes to the wage distribution. After controlling for the changing occupational content of employee wages, the role of average firm residual differences is approximately zero; the modestly increasing trend in between-firm wage inequality is explained by a combination of changes in between-occupation inequality and the occupational specialisation of firms. It is possible that previous studies, which assign some of the importance of changes in the between-firm component to industry, have misrepresented a significant role for occupations. These results are robust across measures of hourly, weekly and annual wages.

Keywords: wage inequality, within-firm inequality, occupational wage premiums
JEL codes: E24, J31

1 Introduction

The long-term trend of rising wage inequality in Great Britain has been extensively documented (Hills et al., 2010; Machin, 2011; Belfield et al., 2017). As in the US and several other countries, the majority of this increase in Britain occurred in the 1980s, but stagnant median real wages in the last two decades have re-focused attention on where the proceeds of growth are ending up. Although well studied, some ambiguity remains over what is the predominant driver of changes in the wage distributions of labour markets such as Britain's. One suggested explanation points towards pay setting practices and the increasingly generous remuneration of executives and senior managers (Piketty, 2013). Others have identified rising skill and occupational wage premiums, potentially driven by skill-biased technological change (Katz & Murphy, 1992; Machin & van Reenen, 1998). Further explanations highlight changing institutions, with major examples in Britain being the decline in unionisation (Card et al., 2004) and the introduction of a minimum wage in 1999 (Machin, 2011). One way to potentially disentangle these explanations is to ask how much have differences between firms, relative to changes within firms, driven recent inequality trends. We attempt to answer this question for the last two decades in Great Britain. The answer matters for at least two major reasons. First, if inequality is rising within firms, it has implications for perceptions of fairness and worker morale, and their theoretical links to productivity growth (Akerlof & Yellen, 1990). Second, it potentially has political connotations, with regards to how salient overall inequality trends are to individual workers. We hope that the results here can help direct future efforts to identify the specific determinants of long-run wage inequality changes, both in Britain and elsewhere.

This paper relates closely to a large and recent literature studying the importance of the firm in determining within country wage inequality trends and patterns. The majority of these studies have found that trends in the overall variance of wages are strongly driven by between-firm (or establishment) differences: by the variance of average firm wages, as opposed to the increasing or decreasing dispersion within firms.[1] Specifically for Britain, Faggio et al. (2010) find that rising wage inequality in the fifteen years to 1999 is almost entirely accounted

---

[1] See amongst others for the US: Davis & Haltiwanger (1991); Dunne et al. (2004); Barth et al. (2016); Song et al. (2016). For Sweden: Nordström Skans et al. (2009); Akerman et al. (2013). For West Germany: Card et al. (2013). For Brazil: Alvarez et al. (2016); Benguria (2015); Helpman et al. (2017). Also, see the recent survey by Card et al. (2016).

for by an estimate of between-firm variance. Prominently, Song et al. (2016) note that the substantial increase in wage variance between US firms has accompanied greater assortative matching of workers and firms. They suggest that the increasing outsourcing of tasks and occupational concentration of firms could account for some part of these results. However, due to a lack of occupational data this has been largely untested.[2]

The main contribution of this paper is to extend the results of Faggio et al. (2010), using the same survey data of wages and hours, but by instead matching a representative sample of employees to the majority of large British firms, for the more recent period between 1996 and 2015. This provides us with a more robust sample of jobs, as opposed to using some separate source to estimate firm average wages; i.e. Faggio et al. (2010) lacked data on wages within firms. The main limitation of our data is that on average we only observe one percent of the employee wages in any firm, unlike the near census data of other recent studies. Nonetheless the data also offer some advantages. They are generally considered to be accurate records from firms' payrolls of annual and weekly earnings, and their constituent components, including hours worked.[3] The dataset contains a detailed classification of occupations, which will prove to be important here, not least given the significant role of the polarisation of work in recent British inequality trends (Goos & Manning, 2007), and the increasing occupational specialisation of firms (Cortes & Salvatori, 2016).

To preview our main results, we find no compelling evidence that overall inequality trends over the past two decades have been mostly driven by changes in firm level differences. This result is consistent across hourly, weekly and annual wage rates. In particular, controls for the occupational content of wages, over time, make firm level differences redundant in accounting for the dynamics of inequality throughout the wage distribution. Some combination of changes to between-occupation inequality and the sorting of occupations across firms must therefore account for any observable role for between-firm differences. This

---

[2]A notable exception is Weber-Handwerker & Spletzer (2016), who do make some progress on this for the US, finding a small part of wage inequality growth between establishments can be explained by the changing extent of occupational concentration. For West Germany, Card et al. (2013) find that intrinsically high wage workers have become increasingly more likely to work for high wage firms, and vice versa for low wages. This effect can explain the vast majority of estimated increasing between-occupation inequality in recent decades.

[3]See for example, Nickell & Quintini (2003); Devereux & Hart (2006); Blundell et al. (2014); Elsby et al. (2016), who due to this perceived accuracy, have used the survey extensively to document the extent of real and nominal wage rigidity over the business cycle.

not only highlights that the drivers of recent wage inequality changes in Britain could be different, but also that the estimated importance of between-firm inequality found elsewhere could similarly disguise an important role for the occupational transformation of firms and labour markets.

The remainder of the paper proceeds as follows: Section 2 describes the data, Section 3 presents the results from decompositions of overall and residual wage variance over the last two decades in Britain, Section 4 describes the dynamics of inequality throughout the wage distribution, and Section 5 concludes. Further information concerning the data, sample construction, mathematical details and additional results are presented in the Appendix.

2 Data

The data we use are from the New Earnings Survey Panel Dataset (NESPD), 1975-2015, which is distributed under secure license access by the UK Data Service, with the permission of the data owners, the Office for National Statistics (ONS). It is a continuing sample of approximately one percent of all Pay As You Earn (PAYE) taxpayers in Britain, with the sample selected using the same last two digits of worker National Insurance numbers each year, covering up to 180 thousand employee jobs per year.[4] A small number of jobs not registered for PAYE, which tend to be of very low pay, are not sampled. Employees who are not paid in the reference period are also excluded. These are both potential sources of composition bias in measuring inequality changes, which could especially vary over the economic cycle. But it is certainly an advantage that the data is a long-running panel, since we can expect many repeated observations of employer-employee matches.[5] Data is collected via a questionnaire issued to employers, who are required by law to respond, and it is intended to be completed with reference to payrolls. They return the gross weekly earnings and hours worked of employees, and their constituent components, as well as an employee's occupation and other information potentially related to remuneration, such as pensions and collective agreements.

---

[4]National Insurance numbers are issued to all individuals in the UK who have the right to work. For UK nationals these are typically issued when turning sixteen.

[5]Though we cannot exploit this feature fully since in the current publicly available form of the dataset employers cannot be identified completely or robustly over time, but only within year.

The reference period for the survey is always a week in April. Gross annual earnings for the preceding year to April have been recorded since 1999.

It is a significant advantage of these data that we can consider the robustness of results across different frequencies of pay. For example, the compositional differences in two jobs samples from the NESPD which contain either non-missing observations of weekly or annual wages could be large, given that for the latter individuals must have been with the same employer for at least twelve months. From 1996 information on employer size and industry classification are added by the ONS from the Inter-Departmental Business Register (IDBR), an administrative census of all UK registered companies.[6] Only very small businesses consisting of the self-employed are not found on the IDBR. The employer reporting unit observed in the NESPD is generally the enterprise or a local unit thereof. For the vast majority of the data used in the analysis that follows the 'firm' is an enterprise.[7] For a sub-period 2004-15, the enterprise of all jobs is identified, including for those whose data were returned at the local unit level, and we use this as a robustness check of whether our less precise definition of a firm could significantly affect any results.[8] This earnings survey is generally considered to be unusually accurate, at least as compared with household based surveys. The NESPD has undergone several minor methodological changes over its lifetime, but the principal aim of collecting detailed and precise information on hours, pay and occupations has remained consistent. In Appendix A we briefly summarise the relevant changes, as well as providing greater detail than what follows on how we construct our analysis sub-samples.

2.1 Creating a large firm sample of the NESPD

For all sub-samples of the NESPD we include only those aged 16-64 and exclude jobs where pay in the reference week has been affected by absence or leave. For weekly wages and hours worked we use reported values excluding any overtime. We also drop a tiny number of jobs

---

[6]This register is compiled by Her Majesty's Revenue and Customs and contains all firms whose turnover is above the Value Added Tax threshold and/or has at least one member of staff registered for income tax collection.

[7]We are comfortable that the enterprise is a typical definition of the firm, as defined for UK government administrative purposes. IDBR definition: "An Enterprise can be defined as the smallest combination of legal units (generally based on VAT and/or PAYE records) that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit."

[8]We can do this using the annual cross-section datasets of the Annual Survey of Hours and Earnings (see ONS), from which the NESPD in later periods is derived by the ONS.

with records of over a hundred hours worked in the reference week. Hourly rates of pay, as in the published ONS summary descriptives, are inferred from gross weekly pay and basic, or usual, hours worked. From 1999 we drop less than half a percent of all observations in each year whose computed hourly rate of pay is less than eighty percent of the applicable National Minimum Wage. When considering annual earnings records we use only observations where the employee has been with the firm for at least a year. All monetary values are deflated to 1997 prices using the ONS' Retail Price Index from April, to match the reference period of the NESPD.[9] To analyse and estimate a within-firm component of wage dispersion we have to match sufficient numbers of employees to each observed firm. Hence we construct a large firm sample of the NESPD. We consider only jobs in each year at enterprises with 250 employees or more, according to the IDBR.[10] In the baseline sample we consider only full-time jobs, defined as working over thirty hours in a week before overtime, and in each year then retain only firms for which there are ten or more job observations with non-missing values of weekly pay and basic hours worked. We construct several other NESPD sub-samples, which are discussed in the results, where we vary the minimum number of job observations per firm and consider not only full-time workers.[11]

Throughout the following analysis and results one can generally replace any reference to 'firms' with 'large firms,' or even 'very large firms.' This is clear when we compare the enterprise size distributions in 2013 of the UK population and the firms in our baseline NESPD sub-sample (Table A1). Over seventy percent of UK enterprises with over 250 employees have less than a thousand employees. But in our baseline sample, such firms are only five percent of the total number. On the other hand, firms with more than two thousand employees are relatively over represented; the sample includes a similar number of firms with over five thousand employees as there are such UK enterprises.[12] Thus, we cannot represent the

---

[9]Accessed from the ONS website 25/05/2016.

[10]The cut-off between the definition of Small and Medium Enterprises (SMEs) and Large firms in the UK is typically at 250 employees.

[11]The restriction of there being at least ten job observations for any firm included in the NESPD sub-sample imposes a de facto minimum size of more than a thousand employees. For the analysis of annual wages the sub-sample contains only firms who have at least ten full-time job observations which have non-missing values for annual pay, and which are at least a year old.

[12]Part of the non-sampling discrepancy here is due to the NESPD being British as opposed to UK. In 2013 ONS data suggests there were thirty enterprises in Northern Ireland with over a thousand employees. Using actual enterprise identifiers from the ASHE to define firms in 2013, and otherwise the same criteria to construct the baseline large firm sample, gives us 598 enterprises with over five thousand employees.

whole firm size distribution of Britain, we can nonetheless claim to sample employees from practically all very large enterprises. As such we are able to study a significant fraction of jobs and wages; for example, in 2013 the firms in our sample represent approximately forty percent of employee jobs.[13]

Given we are studying the dynamics of wage inequality, we also briefly document how the baseline sample's firm size distribution has evolved over time, between 1997 and 2007 for example (Table A2). The share of firms with more than two thousand employees increased by over thirteen percentage points in this period, with the largest increase amongst those with five thousand or more. The share of employee observations in very large firms similarly increased. Despite this, the true distribution of these firms was relatively unchanged over the period, according to their administrative IDBR enterprise level of employment. We believe this reflects a shift since 2004 in the employer reporting unit of the earnings survey towards more commonly being the enterprise, as opposed to the local unit.[14] To give further insight we describe the sample's changing industrial make-up over the same ten years (Figure C1).[15] The share of firms associated with the manufacturing sector decreases notably, whilst real estate and business services firms are increasingly represented, reflecting well-known recent patterns of structural change in Britain. We observe similar trends if we consider the represented labour shares of sectors, though in this case there is also a sizeable decline in the share of employees in public administration and defence.

An advantage of our data over those used in similar studies is the inclusion of employer descriptions of jobs and their assignment to a detailed occupational classification.[16] Comparing the incidence of major occupation groups in the sample over time, some occupations are less prevalent in 2007 than in 1997, with a particularly large decrease for professionals (Table A3). At the same time the share of elementary occupations has increased

---

[13]According to the ONS Labour Market Statistics Workforce Jobs series, there were approximately 27.5 million employee jobs in Great Britain in 2013.

[14]This coincides with the replacement of the NESPD with the ASHE. Despite studying the documentation we cannot find any noteworthy reason for such a sizeable shift.

[15]Throughout the paper industry sectors refers to the Standard Industrial Classification (SIC) 2003, unless stated otherwise. See Appendix A for details on where and how SIC 2007 codes in the NESPD have been cross-walked to SIC 2003.

[16]Throughout, occupations refer to the International Standard Classification of Occupations 1988 (ISCO88), unless stated otherwise. See Appendix A for details on how ONS Standard Occupation Classifications (SOC) were cross-walked.

by almost the same amount. This sample of large British firms, and a further sample of their employees, would appear to have some different characteristics over time. Partially this could reflect well-known long-run trends of structural change, but could also be a feature of how we construct the sample and focus on only the part of the economy populated by large firms in any period.

## 2.2 Describing wages in large firms and the NESPD

Since economy-wide trends in wages have been extensively documented elsewhere using the NESPD (e.g. Machin, 2011), here we only focus on whether recent patterns amongst jobs in large firms have been notably different. Figure 1 compares selected percentiles of real log wages for full-time employees between our baseline sub-sample and the whole NESPD. Figure C2 similarly compares mean values. All measures of real wages were relatively stagnant during the 2000s. They also experienced a substantial decline since 2008, especially as compared with other downturns. The variance of log weekly wages increased for the whole NESPD persistently from 1975 to 1995 (Figures 1a). The variance of wages in our baseline sample is somewhat lower than in the whole NESPD. This is due to a tighter distribution of wages above the median amongst those working in larger firms. Generally though the pattern of wages across the large firms distribution is similar to the whole NESPD: for example, both show a steep increase in hourly and weekly wages for top earners in the early 2000s, as well as a decline in variance at the onset of the Great Recession, driven by relatively higher earnings at the bottom. Figure 2a further demonstrates these changes by plotting wages relative to 1996 for selected percentiles of the large firm sample. The increase in the variance of log annual wages, which include performance related payments, is substantial between 1999 and 2007 (see also Figure 3b). As shown in Figures 1b and 2b, this is explained by real wages falling at the lower percentiles and only marginally rising at the median, whilst the ninetieth percentile increased consistently through this period. Much of the increased variance during the preceding decade was reversed in 2008 by a relatively greater increase in log wages at the bottom, a compositional effect of the Great Recession. These patterns are similar when we consider all employees and not only those working full-time (Figures C3-C5).

FIGURE 1: Percentiles of real log wages in large firms, full-time employees only, and comparison with the whole NESPD sample, 1975-2015
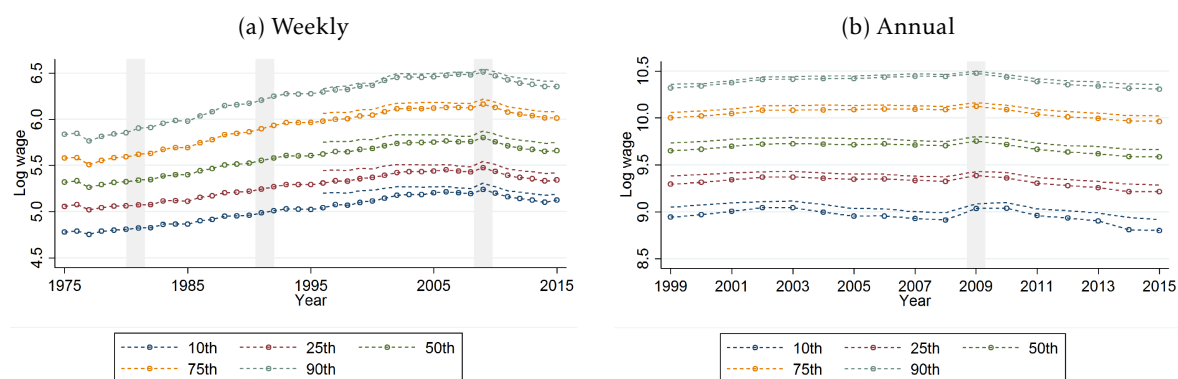
(a) Weekly

(b) Annual

Notes.- author calculations using the NESPD, age 16-64 and full-time employees only. 'Weekly' exclude overtime. 'Annual' are for employees with the firm at least one year. See the text for further details of sample construction. Shaded areas represent official UK recessions. Dashed lines without markers are the series for a large firm sub-sample of the NESPD.

FIGURE 2: Percentiles of real log wages in large firms, full-time employees only: differences relative to 1996/9

(a) Weekly

(b) Annual
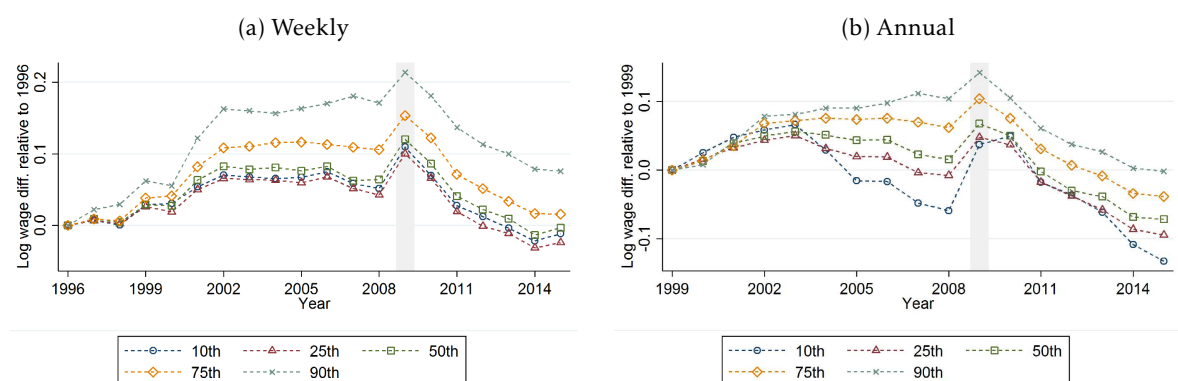
Notes.- see Figure 1.

3 Wage inequality trends: the role of between-firm variance

To account for how much of the variance in employee wages is explained by differences in the average wages paid by firms, we use the well-known decomposition of Davis & Haltiwanger (1991). This approach is used widely in the related literature. The total variance of the natural logarithm of wages across a set firms and their employees V_e can be decomposed into a within-firm component V_{wf} and the variance of average log wages between firms V_{bf}. From our samples we estimate this decomposition as follows. Denoting the total number of firms in a given year by J, and the number of employees we observe in firm $j = 1,...,J$ by N_j,

8

such that the total sample number of employees is $N = \sum_{j=1}^{J} N_j$, then we can write

$$\underbrace{\frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{N_j}(w_{ij}-\bar{w})^2}_{\text{Overall - }V_e} = \underbrace{\frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{N_j}(w_{ij}-\bar{w}_j)^2}_{\text{Within-firm - }V_{wf}} + \underbrace{\sum_{j=1}^{J}\frac{N_j}{N}(\bar{w}_j-\bar{w})^2}_{\text{Between-firm - }V_{bf}},\qquad(1)$$

whereby w_{ij}, $\bar{w}$, and $\bar{w}_j$ denote respectively the log wage of employee i in firm j, the sample mean of log wages, and the sample mean of log wages within firm j.[17] The term capturing the between-firm component of wage dispersion weights by employment share the observed distance of a firm's estimated average wage to the overall average wage, such that larger firms have a potentially greater influence on wage dispersion than smaller firms.[18]

Table 1 summarises the decomposition results discussed throughout this section. Since the data are not top-coded, we exclude the top one percent of earners from all calculations in this section. Throughout the remainder of the paper we mostly focus on weekly wages, as these are recorded in the data independently of an employer's response for the hours worked of their employees. Further, this sample includes jobs which are less than a year old. These jobs would be excluded from an analysis of annual wages and their importance within the true wage distribution could vary over time. Figure 3 plots the components of (1) for each year between 1996 and 2015 for full-time employees in our sample. Total wage dispersion is increasing when measured over the entire sample period (column (9), Table 1). However there is an observable difference pre and post the 2008 financial crisis. The latter period experienced falling inequality, mostly accounted for by the decreasing variance of wages within firms, whilst at the same time between-firm inequality continued to increase. Prior to 2008, the increase in within-firm inequality explained the majority of the overall trend (over 80 percent: column (7), Table 1). The overall variance of log weekly wages mirrors closely the

---

[17]Sampling errors in the values of firm average wages (or hours) will positively bias between-firm variance estimates and their shares of the overall variance. We do not attempt to correct this, and instead rely on our analysis being focused on trends, since the size of this bias is unlikely to vary significantly over the period studied. The literature in this area, such as Card et al. (2013), also acknowledges the bias from sampling error, and similarly tends to ignore it, by arguing that trends are unlikely to be affected. Here we are especially reliant on any changes to the NESPD/ASHE sample frame or method not affecting the level of bias over time. We are confident that this is qualitatively the case, given our knowledge of the timing of any such changes, as discussed in Appendix A.

[18]Within the data there are two potential choices for how to weight firms: by their share of employee observations in the sample, or their relative size as indicated by their IDBR recorded number of employees. Our preference throughout the paper is the former, but we find it has no qualitative effect on results throughout. We demonstrate this for an example in Figure C6.

pattern of the within-firm component, as can be seen in Figure 3a.[19] It is clear that the pre 2008 increase and the post 2008 decrease in inequality are driven mostly by the within-firm component in Britain, which contrasts starkly with previous findings for other countries. A similar conclusion holds for annual wages in Figure 3b, which exhibit more substantial inequality changes over the period, most likely explained by the inclusion of overtime and bonus payments (see also Figure C9).[20] Just over forty percent of the long-run increase in annual wage inequality is accounted for by between-firm variation, compared, for example, with sixty percent found by Song et al. (2016) for jobs in large US firms. A decrease in wage dispersion during the financial crisis is also more pronounced in annual wages, and accounted for mostly within firms, perhaps reflecting again the presence of performance related and bonus pay.

FIGURE 3: Within and between-firm components of the variance in log employee wages, 1996-2015
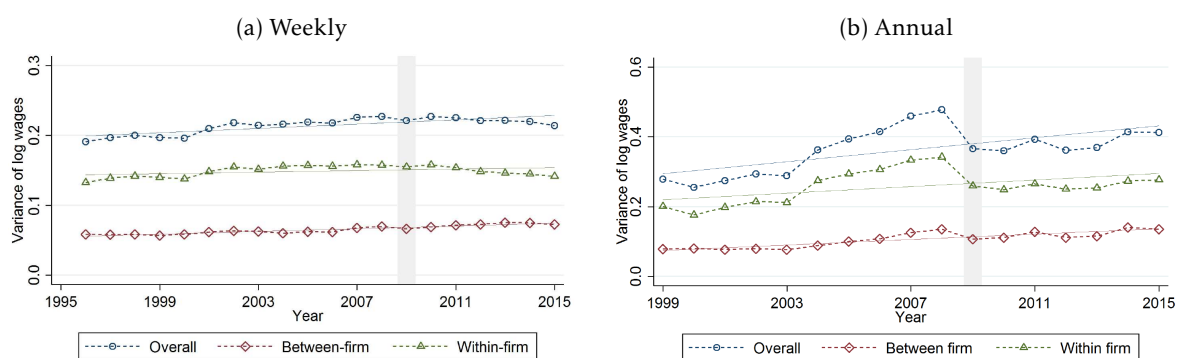
Notes.- author calculations using the NESPD, age 16-64 and full-time employees only. 'Weekly' exclude overtime. 'Annual' wages are for employees with the firm at least one year. The data is for all large firms in the NESPD who have at least ten full-time employee wage observations in that year. The top one percent of wage values in each year are excluded from calculations here. Shaded areas represent official UK recessions. Lines without markers are linear trends.

It is apparent that over the last two decades any short or medium-term inequality changes are not driven by the between-firm component. Overall wage inequality exhibits stronger co-movement with its within-firm component than the between, implying that the latter is less important in driving any overall changes. This result also holds when we consider

---

[19]The two series have a correlation coefficient of 0.85, compared to 0.7 for the overall and between-firm component.

[20]Barth et al. (2016) show that conditioning the sample on "stayers", as we have done here, relatively dampens the measured change in US annual wage inequality. However, the relevance of that result here is limited since we are instead comparing across different frequencies of pay and possibly quite different sample compositions of workers.

three sub-samples, each consisting of approximately a third of the employee observations: the public sector, SIC 2003 sectors G-H (wholesale, retail, hotels restaurants etc.) and the remainder of the private sector (Figure C7). Given we consider only full-time employees at this point, unsurprisingly the results are qualitatively unchanged for hourly wage inequality (see Figure C8). Where we can identify firms exactly at the enterprise level, using the ASHE datasets for 2004-2015, results are also not qualitatively different (Figure C10).

It is a substantial advantage of our data that we can confidently identify the determinants of actual wage inequality as opposed to earnings, by being able to further decompose weekly wages into components which account for the variance in log hourly wages and hours worked, and their covariance:

$$V_{wf} = V_{wf}^w + V_{wf}^h + 2Cov_{wf}(w,h), \tag{2}$$

$$V_{bf} = V_{bf}^w + V_{bf}^h + 2Cov_{bf}(w,h) \tag{3}$$

(see Appendix B.1 for exact definitions and derivations of these terms). The covariance terms are potentially significant, since both individual and firm average wages and hours are known to be strongly correlated.[21]

Figure 4 plots the decomposition described by (2)-(3) for weekly wages. Unlike other related studies, we can show explicitly that the variation in hours worked does not affect the results for full-time employees. Both between and within-firm hours variance components together account for less than five percent of weekly wage variance throughout the period (column (2), Table 1). This provides some support then to results in other studies which cannot directly observe hours, but restrict their attention to full-time employees, such as in Card et al. (2013). When we contrast this with a decomposition of weekly wage variance which includes those working part-time, changes in the variance of hours worked within firms most closely determine overall inequality changes; in the last two decades firms have been increasingly using of a mix of part and full-time employees. However, the sharp increase in wage variance

---

[21]This advantage of our data is emphasised in a recent study by Belfield et al. (2017). Using potentially less accurate household survey data for all employees, they find that a sixth of the increase in male log weekly earnings variance over the past two decades in Britain is accounted for by greater hours variation. The increased tendency of low wage work to accompany low hours accounts for a further thirty percent. They also find that the entire fall in female wage inequality is explained by these factors, and not by any change in wage rate variance.

amongst all employees in 2004-05 observed here coincides with a methodological shift in the survey, whereby more low-paid and part-time jobs without PAYE numbers were sampled. We discuss this further in the Appendix, but it is a good reason why we mostly focus on only full-time workers in the analysis here. In terms of levels, hours components account for as much as forty percent of overall wage inequality. The covariance in hours and wages, both within and between firms, is also a significant component, together accounting for as much as twenty percent, reflecting the tendency of part-time jobs to be more commonly low-wage.

FIGURE 4: Within and between-firm, hourly rate and usual hours components of the variance in log weekly employee wages, 1996-2015

(a) Full-time

(b) All employees
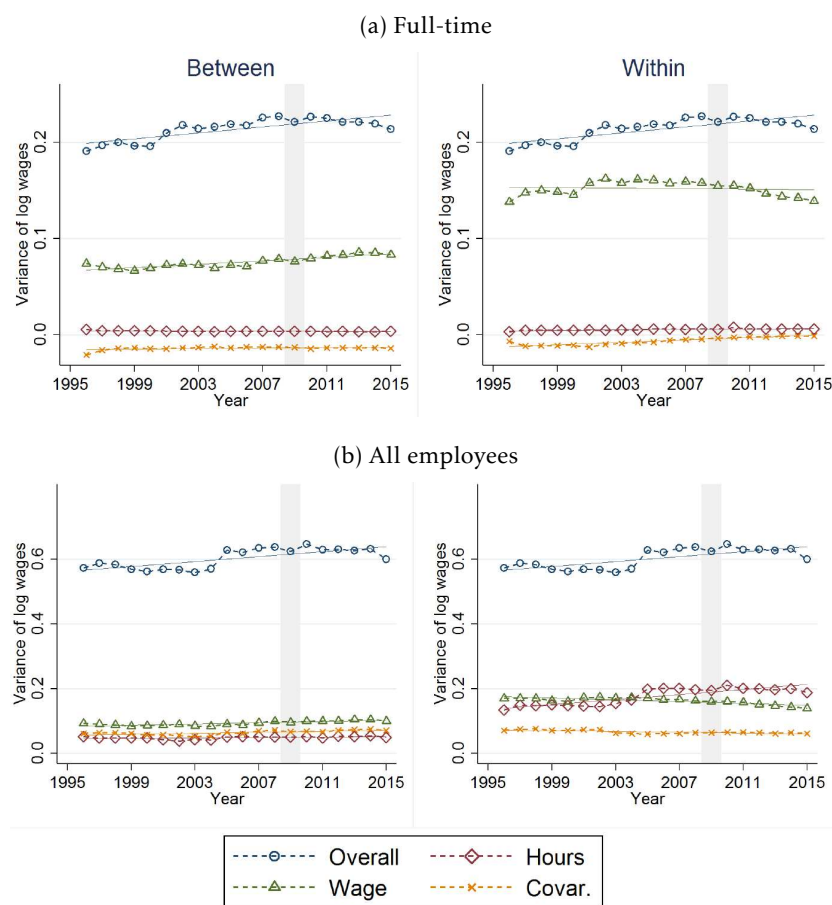
Notes.- author calculations using the NESPD, age 16-64 only. Wages and hours exclude overtime. The top one percent of wage values in each year are excluded from calculations here. The data is for all large firms in the NESPD who have at least ten (full-time) employee observations in that year. The 'Covar.' series represent twice sample covariance terms. The "Overall" series, in both left and right panels, is the total sample variance. As such, all other series across both panels sum within year to this total variance. See the text for further details of how the sample is constructed. Shaded areas represent official UK recessions. Lines without markers are linear trends.

TABLE 1: Summary of decomposition results for the variance in log employee wages

	1996 Level (1)	1996 Share (2)	2005 Level (3)	2005 Share (4)	2015 Level (5)	2015 Share (6)	Change 1996-2005 Level (7)	Change 1996-2005 Share (8)	Change 1996-2015 Level (9)	Change 1996-2015 Share (10)
Raw weekly wages:	0.191		0.219		0.214		0.028		0.023	
Between: total	0.058	0.305	0.062	0.284	0.072	0.339	0.004		0.014	
Hourly wages bf	0.074	0.387	0.072	0.331	0.083	0.389	-0.001	-0.056	0.009	0.002
Usual hours bf	0.005	0.029	0.004	0.016	0.004	0.016	-0.002	-0.012	-0.002	-0.012
Covariance bf	-0.021	-0.110	-0.014	-0.063	-0.014	-0.067	0.007	0.047	0.007	0.043
Within: total	0.133	0.695	0.157	0.716	0.141	0.661	0.024	0.021	0.009	-0.034
Hourly wages wf	0.138	0.723	0.161	0.734	0.139	0.65	0.023	0.011	0.001	-0.073
Usual hours wf	0.003	0.016	0.006	0.027	0.006	0.028	0.003	0.011	0.003	0.012
Covariance wf	-0.008	-0.044	-0.01	-0.045	-0.003	-0.016	-0.001	-0.001	0.005	0.027
Resid. weekly wages:	0.088		0.101				0.013			
Between	0.022	0.246	0.020	0.198			-0.002			
Within	0.066	0.754	0.081	0.802			0.014	0.049		
Raw annual earnings:	0.279		0.393		0.412		0.115		0.133	
Between	0.078	0.281	0.010	0.254	0.135	0.328	0.022		0.057	
Within	0.201	0.719	0.293	0.746	0.277	0.672	0.093	0.027	0.076	-0.047
Resid. annual earnings:	0.214		0.317		0.412		0.103		0.198	
Between	0.037	0.174	0.049	0.154	0.135	0.328	0.012		0.098	
Within	0.177	0.826	0.268	0.846	0.277	0.672	0.092	0.020	0.100	-0.154
Raw weekly wages:	0.574		0.629		0.601		0.055		0.027	
Between: total	0.204	0.356	0.204	0.325	0.218	0.363	0.000		0.014	
Hourly wages bf	0.092	0.160	0.090	0.143	0.099	0.165	-0.002	-0.018	0.007	0.005
Usual hours bf	0.051	0.090	0.050	0.080	0.049	0.082	-0.001	-0.009	-0.002	-0.008
Covariance bf	0.061	0.106	0.064	0.102	0.07	0.116	0.003	-0.004	0.009	0.010
Within: total	0.370	0.644	0.424	0.675	0.383	0.637	0.055	0.031	0.013	-0.007
Hourly wages wf	0.169	0.295	0.171	0.272	0.139	0.231	0.002	-0.023	-0.031	-0.064
Usual hours wf	0.133	0.232	0.199	0.316	0.188	0.312	0.066	0.085	0.055	0.081
Covariance wf	0.068	0.118	0.055	0.087	0.056	0.094	-0.013	-0.03	-0.011	-0.024

Full-time employees *(rows: Raw weekly wages through Resid. annual earnings)*

All employees *(rows: Raw weekly wages through Covariance wf)*

Notes.- author calculations using the NESPD. See text for further description of the data sample and methods. *bf* and *wf* refer to the between and within-firm components of the overall variance. Relevant rows may not sum exactly due to rounding.

3.1 Residual wage inequality

In Section 3 we described how the baseline sample has changed over time in terms of firm size, and the industry sectors and occupations represented. Further, we cannot be certain that our results are not affected by changes in how much some characteristics of jobs are rewarded: for example, the recent rise in the London wage premium, which could potentially manifest as greater between-firm inequality.[22] To account for this, we regress log wages in each year of our sub-samples of the NESPD on sets of employee and employer characteristics, and then describe inequality in the resulting residuals. In each wages equation we include a minimum set of controls for sex, age and its square, and the region of employment. We also consider controls for industry sectors and occupations at varying levels of detail, as well as the IDBR enterprise size of the firm.[23] All remaining heterogeneity in wages, originating from employee, firm and employee-firm specific premiums, is left in the error term.

Figure 5 compares the estimated residual log weekly and annual wage variance with the equivalent overall non-residual variance, using alternative specifications of the wages regression described above. For both measures of wages, the patterns over time appear to be mostly unaffected by the inclusion of controls for regions, age groups, gender and industry sectors. This implies that any dynamic changes in the overall composition of our baseline sample of the NESPD and/or wage premiums for observable characteristics are insignificant. In other words, the changes in actual British wage inequality over this period are mostly residual changes, occurring within regions, age groups, gender and industry sectors. Replacing the controls for industries with occupations however does lead to less pronounced inequality dynamics in the estimated residuals. Therefore, changes in between-occupation inequality and the sorting of workers across occupations must be significant factors in the last two decades. The level of weekly wage variance is reduced by approximately a half with controls for sex,

---

[22]For example, ONS published results from the ASHE for the nominal median weekly pay of full-time employees shows an increase between 1997 and 2007 of forty-five percent in London, compared with thirty-five percent in the North East. Another notable trend in British wage premiums is the closing of the Gender Pay Gap over that time period.

[23]We do not have information on years of education, or some other explicit proxy for levels of human capital. The only way we can mitigate the resulting concern, that this missing information would be correlated with occupation controls, is by considering the robustness of any results whilst varying the detail of the occupational classification used.

age, region and occupation groups.[24] This fraction does not increase substantially when we consider a more detailed classification of occupation groups or add industry controls. Thus, these broad occupation groups not only explain a large portion of the wage variance, but also the inequality of residual wages is not greatly affected by conditioning on a more detailed classification of jobs.

FIGURE 5: Variance of residual log employee wages, 1996-2015
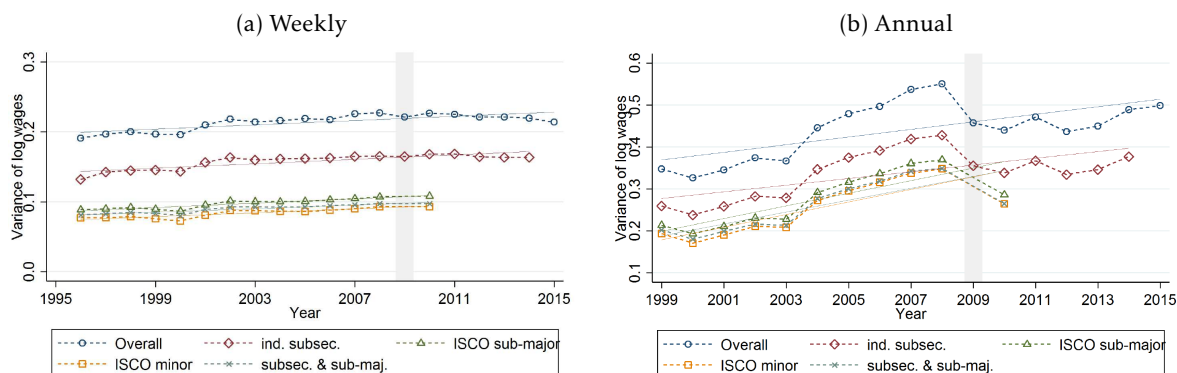
(a) Weekly	(b) Annual

Notes.- author calculations using the NESPD, age 16-64 and full-time employees only. 'Weekly' exclude overtime. 'Annual' are for employees with the firm at least one year. The top one percent of wage values in each year are excluded from calculations here. The data is for all firms in the NESPD who have at least ten full-time employee observations. "Overall" gives the non-residual variance. All residual log wages are calculated using OLS with controls for sex, age, age squared and major regions. Also presented here are results with controls in addition for SIC 2003 2 digit, ISCO sub-major groups, ISCO minor groups, & SIC 2003 2 digit with ISCO sub-major groups. Shaded areas represent official UK recessions. Lines without markers are linear trends.

We can also account for the role of between-firm differences in residual wage inequality changes. Figure 6 shows that by conditioning on industry, occupation or both, the share of residual wage variance accounted for by the within-firm component rises substantially. Unsurprisingly, a large part of the difference in average wages across firms is accounted for by their industries and the types of workers they employ. For residual weekly wages especially, with controls for occupations the within-firm share increases over time relative to the equivalent measure for actual wages. More importantly for the discussion here, there is no suggestion that the between firm component is driving the overall pattern of residual wage inequality.

---

[24]Although the total number of sub-major occupation groups controlled for in the sample is twenty-nine, in practice, nine of the groups constitute less than two percent of observations across all years.

FIGURE 6: Share of within-firm component in the variance of residual log employee wages, 1996-2015

(a) Weekly

(b) Annual
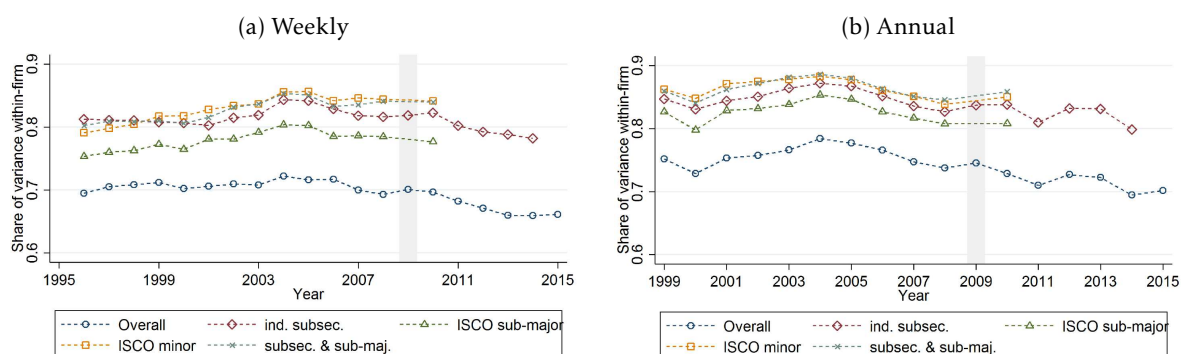
Notes.- see Figure 5.

4 Inequality changes throughout the wage distribution

In analysing the dynamics and components of an aggregate measure of wage inequality we could miss a more complex evolution of the cross-sectional wage distribution over time. To determine the role of firms in changes across and within the distribution of wages we employ a graphical method of analysis popularised by Juhn et al. (1993), and subsequently adapted by Song et al. (2016) and Benguria (2015) amongst others. Simply note that we can re-write log wages as follows,

$$\underbrace{w_{ij}}_{\text{Employees}} = \underbrace{\bar{w}_j}_{\text{Firms}} + \underbrace{\left[w_{ij} - \bar{w}_j\right]}_{\text{Employee/firm}} . \tag{4}$$

We compute estimates of the averages of each term in (4) within each percentile bin of the employee wage distribution in every year. By considering the resulting differences across percentiles and between years, we can then account for the role of firm average wages, as opposed to the relative difference between employees' wages and their firms' averages, in driving wage inequality changes.

Figure 7 represents this decomposition for the change in real log weekly wages between 1997 and 2007, using the baseline sub-sample of full-time employees. The relatively smooth "Employees" series plots the change in the average log wage of workers in each percentile between the two years. To avoid confusion, these are unlikely to be the same individuals; this is a comparison of annual cross-sections. Each percentile is decomposed using around four to five hundred job observations in each year. A positive slope across percentiles indicates that

16

FIGURE 7: Change 1997-2007 in the average log weekly wage by percentile of employees and the contribution from firms

Notes.- author calculations using NESPD, age 16-64, full-time employees only. Weekly exclude overtime. The data is for all firms in the NESPD who have at least ten full-time employee observations. The "Employees" values are computed by taking the average log real wages of employees within each percentile, increasingly ordered by the level of wages in both years, and taking the difference across years. The "Firms" values are computed by taking the average across workers, in each percentile, of the average log wages of the firms they work for, in each year, and then taking the difference across years. The "Employees/Firms" values are the residual of these other values: equivalently, the average across workers, in each percentile, of the log difference in employee wages from their firms' avearage value, in each year, and taking the difference across years.

in some portion of the wage distribution inequality has increased. For example, wages at the median increased by 0.05 log points over this period, but by 0.1 points at the seventy-fifth percentile, and 0.2 points at the ninety-fifth. Representing the evolution of wage inequality in this way shows that relatively small changes in the time series of overall wage variance can bely starker inequality dynamics. For example, here we can see that inequality has fallen amongst the very lowest earners, undoubtedly in no small part due to the introduction of the National Minimum Wage in 1999. By construction, the average level of the "Firms" components across percentiles is the same as that for "Employees", and the "Employee/firm" component is centred about zero. For the graphical analysis it is the slopes of these series across the percentiles which concern us.

The firms component contributes somewhat to the rise in wage inequality at the top of the wage distribution, but the employee/firm component also contributes, increasing across the percentiles from the twentieth onwards. This is consistent with results for the US in Song et al. (2016), that amongst large firms the between-firm component appears to not be wholly driving wage dynamics. However, in Great Britain for this period, for smaller firms than what

are considered large in Song et al. (2016), the firms component is weaker. Before progressing further, we also represent the change in inequality since the Great Recession, for 2008-2015, in the same way (Figure C11). We can demonstrate here that our results are unaffected if to define firms we instead use the administrative definition of an enterprise from the ASHE datasets. Real wages decreased across the whole distribution since the financial crisis, but inequality also fell. However, there is no suggestion in the data that this can be accounted for by changes in the differences in average wages between firms.
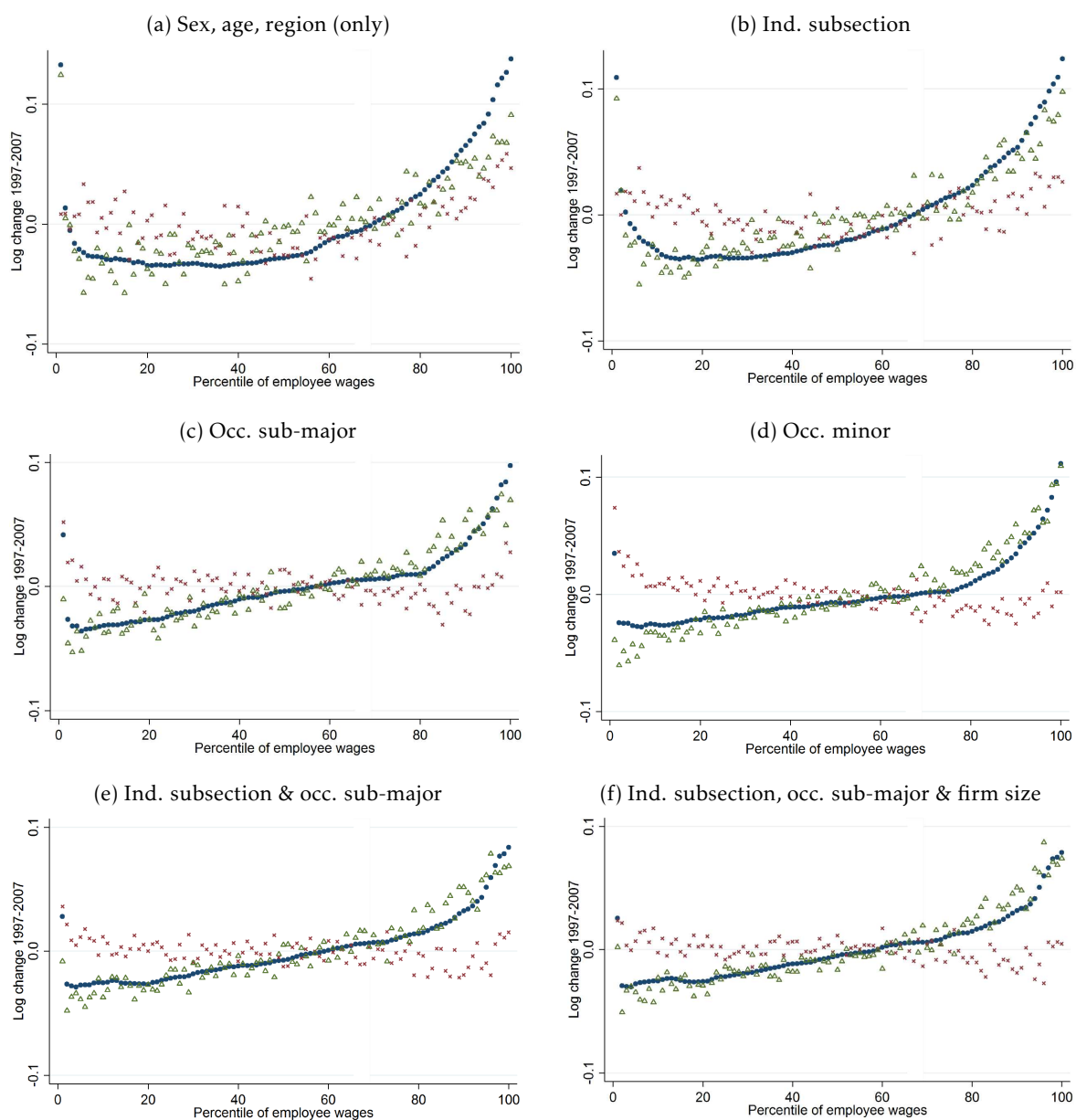
4.1 Residual wage percentiles: the role of occupations

As we have shown above, a substantial fraction of wage variance unsurprisingly is accounted for by some simple employer and employee characteristics. We also showed that accounting for worker heterogeneity and changing wage premiums reduces the amount of inequality which can be attributed to between-firm differences. In Figure 8 we consider the same graphical analysis as above, but using the residual weekly wages of full-time employees. For brevity, we ignore the equivalent analysis for predicted wages since any changes in inequality they account for are much smaller in magnitude.[25]

We present results here using several variants of the wage regression. Including only controls for sex, age and region of work leaves the pattern of wage inequality changes relatively unaffected, when compared with Figure 7, with wages at the top still having moved substantially away from the median, but with little change below the median. The within and between-firm series remain noisy across the percentiles, but there is still no sense that average firm wages explain the majority of the relatively greater rise in wage premiums at the top. After adding industry controls, this becomes even clearer, and the firms contribution more or less disappears. This is somewhat consistent with the variance decomposition of US establishment level average earnings carried out by Barth et al. (2016), who find that up to a half of its increase in the three decades to 2007 can be accounted for by controls for industry specific wages. If we instead control for occupational wage premiums, the pattern of employee wage changes across the percentiles noticeably alters. There is then evidence of rising inequality across the distribution, with the slope becoming steeper from

---

[25]See again Figure 5, or for the US equivalent result which corroborates this see for examples Juhn et al. (1993) & Barth et al. (2016).

FIGURE 8: Change 1997-2007 in the average residual log weekly wage by percentile of employees and the contribution from firms

(a) Sex, age, region (only)

(b) Ind. subsection

(c) Occ. sub-major

(d) Occ. minor

(e) Ind. subsection & occ. sub-major

(f) Ind. subsection, occ. sub-major & firm size

Notes.- author calculations using NESPD, age 16-64, full-time employees only. Weekly exclude overtime. The data is for all firms in the NESPD who have at least ten full-time employee observations. All residual log wages are estimated using OLS with controls, in addition to those stated, for sex, age, age squared, and major regions. See notes for Figure 7 or the text for a description of how series are calculated and interpretation.
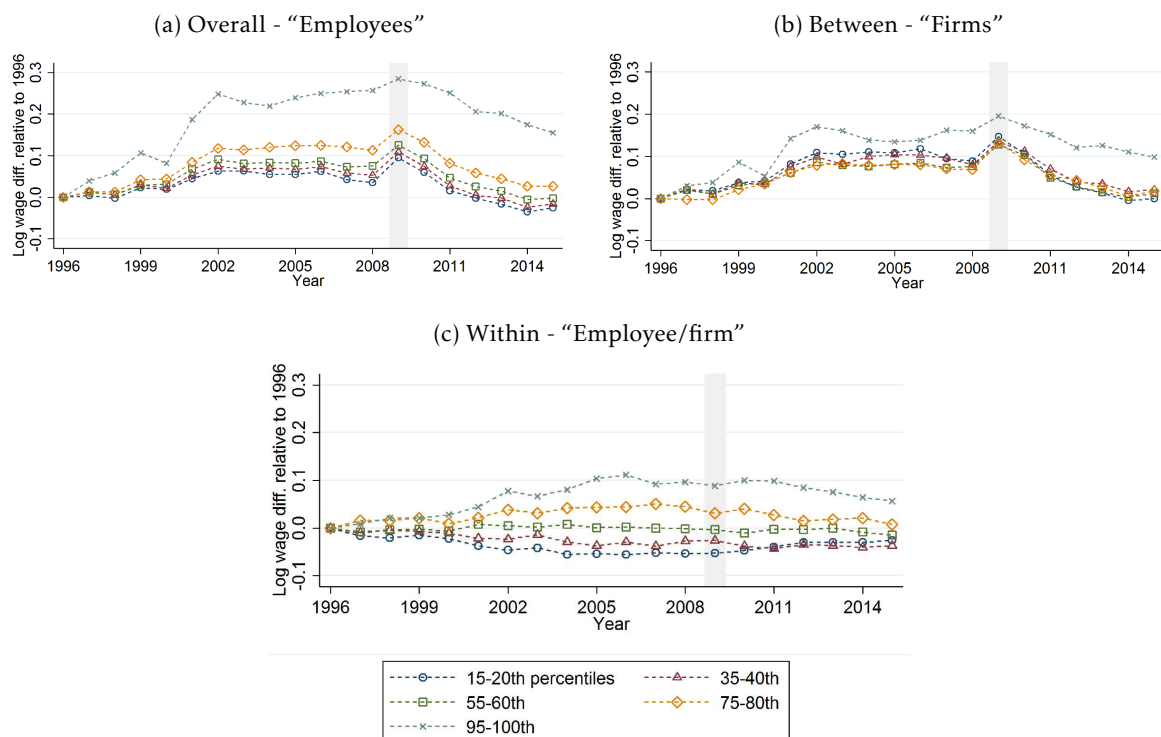
the eightieth percentile upwards. Firm average residual wages did not account for these dynamics: neither for rises below the eightieth percentile nor the greater increase in the highest employee residual wages. This result becomes even clearer when we consider more detailed classifications of jobs, and adding controls for firm size makes no difference. Once we condition on the occupational content of wages, the firm specific component becomes almost

19

irrelevant in accounting for residual wage inequality dynamics. Any change in wage inequality across large firms is due to a combination of changes in between-occupation inequality and the concentration of high or low wage occupations within firms. Further, given that the main result here does not appear to depend on controls for industry, it is possible that previous results in other studies, which assign some of the importance of changes in the between-firm component to industrial change, are to an extent misrepresenting a more significant role of occupations.

4.2 Some robustness checks

We consider how robust these results are to variations in the time period studied. Still focusing on full-time weekly wages, Figure 9 considers the changes for selected ventiles of the non-residual employee weekly wage distribution relative to 1996. The average wages paid by firms can account for some of the relatively greater increase in the top five percent of employee wages in the early 2000s. But dispersion within firms explains most of the overall inequality patterns over the last two decades.

FIGURE 9: Average log weekly wage of employees in selected ventiles, relative to 1996, and contributions from firms

(a) Overall - "Employees"

(b) Between - "Firms"

(c) Within - "Employee/firm"

15-20th percentiles   35-40th
55-60th   75-80th
95-100th

Notes.- see Figure 7.

20

Figure C13 looks at changes over five years across all percentiles for three sub-periods. Notably for robustness, these are periods where the classification of occupations used in the NESPD is constant, and thus cross-walking across classifications was not necessary. As also seen above, the majority of recent increases in employee wage inequality occurred in the five years to 2001. There is no contribution to this from the firms component for residual log wages. This is also the case for non-residual wage inequality, apart from some contribution to greater changes above the ninety-fifth percentile. For 2002-07 and 2005-10, the rise in wage inequality is small, and this is driven by greater wage changes for only the highest earners. Figure 10 replicates Figure 9 but instead for residual wages. There is no substantial contribution from between-firm inequality to the dynamics of the residual wage distribution since 1996. Figure C12 further demonstrates the robustness of this result across all percentiles, considering changes over other ten year periods, each beginning in a year between 1996-2000.

FIGURE 10: Average residual log weekly wage of employees in selected ventiles, relative to 1996, and contributions from firms

(a) Overall - "Employees"

(b) Between - "Firms"

(c) Within - "Employee/firm"
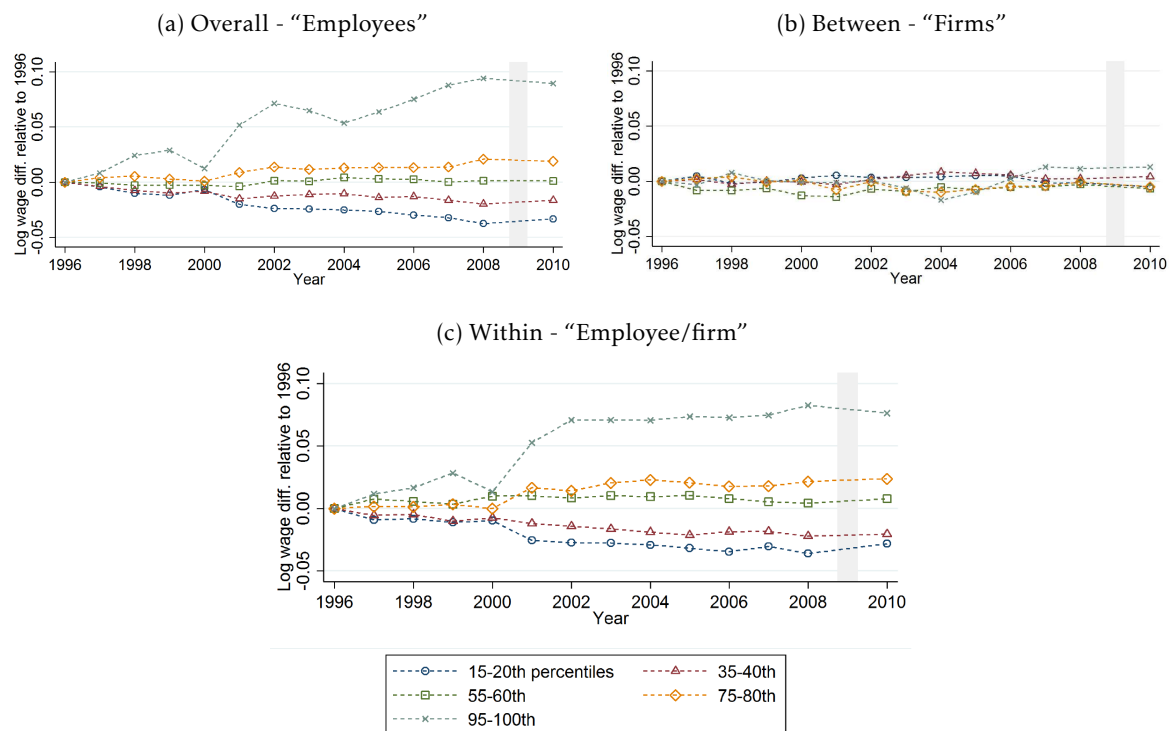
Notes.- see Figure 8.

So far we have only discussed the dynamics of weekly wages for full-time employees working for firms with at least ten job observations in the NESPD in any given year. We

can also check whether results change for the period 1997-2007 when we alter these aspects of the sample. Figures C14-C16 decompose the log change in the weekly wages of full-time employees who are employed by large enterprises with at least one, five or twenty employee job observations. For actual wages, as we increase the sample size and include some smaller firms, it becomes clearer graphically that the firms component cannot explain inequality dynamics. Considering residual wages, with controls for occupations, the results are also qualitatively unchanged as we vary the average firm size in our sample. In Figure C17 we return to our baseline sample, but now study only private sector employees. Again, the results are unaffected. Further, Figure C18 shows that there is no qualitative difference in results if we decompose hourly wage dynamics as opposed to weekly. For annual wages, Figure C19 demonstrates that for non-residual wage inequality all of the dynamics across percentiles are explained by the changing picture within firms. This is also the case when we turn to the inequality in annual wage residuals. Finally, we also consider the picture for weekly wages including part-time employees, and after conditioning on employee characteristics, there is no suggestion in Figure C20 that firm average wages have driven inequality dynamics in this case either.

## 4.3 Reconciling our results with the existing literature

Despite finding evidence for this short period in Great Britain that is contrary to much of the recent literature, we nonetheless believe we can reconcile our results with some of these previous studies of inequality dynamics. First, our analysis is dominated by the very largest firms in Britain. Already Song et al. (2016) have shown that in the US firm size matters.[26] Larger firms come from a starting point of having more diverse workforces and complex pay structures, and so there is far more scope for changes over time in the dispersion of wages within as they evolve. Second, we believe our results chime strongly with a hypothesis from Song et al. (2016): the reason within-firm inequality cannot account for overall dynamics in most studies could be due to the increasing occupational concentration, or specialisation, of firms, coinciding with falling costs of outsourcing work tasks, and a greater tendency to focus on so-called "core-competencies." The very large and long-lived firms which dominate our

---

[26]Similarly, Mueller et al. (2016) have found that inequality pay between job titles within British firms is increasing with their size.

sample are where we might expect such changes in the degree of specialisation to mostly occur. Further, adding to this the continued trend of increasingly polarised demand for occupations in the British labour market, it is then no longer surprising that once we focus on the inequality dynamics of residual wages, which control for changing occupational wage premiums and the composition of the workforce, the role of innate between-firm differences becomes markedly weaker, or even non-existent.

To illustrate the above points further, we focus on the "Firms" component of the change between 1997 and 2007 in weekly wage residuals for full-time employees, represented by Figure 8b: i.e. with controls for industry sectors but not occupations in the log wage regression. Averaging across employee wage percentiles, we carry out a shift-share decomposition of this component. This accounts for the role of the changing occupational structure of the firms represented in each decile (see Appendix B.2 for details). Figure 11 shows part of this decomposition. The "Wages" component, which is computed by holding the occupational structure of firms representing the employees in each decile constant, and allowing only wages to change, does not correlate across percentiles with the overall "Firms" component; the between-firm inequality increase through the top deciles is mostly accounted for by the changing occupational structure of the firms who pay the highest wages, holding occupational wage premiums constant.

FIGURE 11: Decomposing the firm component of employee wage inequality patterns, 1997-2007: the role of changing firm occupation shares vs wages
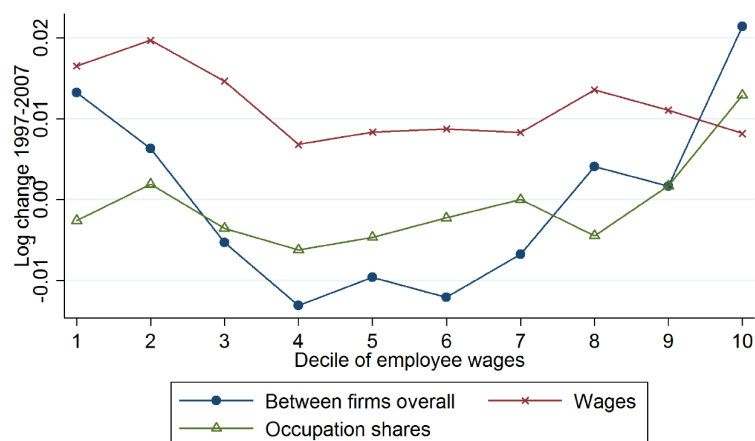
Notes.- this figure takes the average over deciles of the firm component of wage changes, as in Figure 8b (circle markers), and carries out a shift-share decomposition. The two components thereof reported here are as follows: first holding the average across employees of firm occupation shares constant, and considering only average occupational wage changes (cross markers), and second holding the average across employees of firm average occupational wages, in a decile, constant but varying only firm occupation shares (diamond markers).

5 Conclusion

We have used well-known methods to answer whether recent changes in British wage inequality, viewed through a sample of employees at mostly very large firms, can be accounted for by between-firm inequality. We have found substantial evidence that in the last two decades this has not been the case. This is also clear when we consider wage residuals, controlling for changes to occupational premiums and the composition of employment. At first look this would appear to contradict what is becoming a stylised fact, across several countries, that between-firm wage inequality is the most important driver of overall trends. But this is not the first paper to suggest that some part could be accounted for by the changing supply and demand of occupations across firms and labour markets (see Card et al., 2013; Song et al., 2016). The results here strongly suggest that future analyses of this kind should attempt to seek out data which can address the possible role of the changing occupational structure of firms. Otherwise it could be challenging to identify whether inequality changes are accounted for by some unexplainable greater segregation of workers across firms, or whether this to some extent reflects only the combined effects of changes to the occupational polarisation of employment and firm level specialisation. In other words, the role of assortative matching over innate firm and worker productivities could be overstated.

A significant limitation of the analysis here is that we are restricted to studying repeated cross-sectional data of jobs and wages, since employers cannot be identified reliably across time for any extended period in the NESPD. Furthermore the results here only reflect what has happened for the wages in mostly very large firms. We believe this is the limit of what can be achieved using currently available British data sources. We hope that existing UK administrative earnings data, for all employees and their employers, will become available for research in the near future. Only then can the continuing large evidence gap regarding the determinants of British wage inequality be more completely addressed, with the NESPD's more detailed records of job characteristics, such as hours and occupations, serving as a useful supplementary data source.

References

Akerlof, G., & Yellen, J. L. (1990). The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics*, *105*(2), 255–283.

Akerman, A., Helpman, E., Itskhoki, O., Muendler, M.-A., & Redding, S. (2013). Sources of Wage Inequality. *American Economic Review: Papers & Proceedings*, *103*(3), 214–219.

Alvarez, J., Benguria, F., Engbom, N., & Moser, C. (2016). Firms and the Decline of Earnings Inequality in Brazil. Working paper.

Barth, E., Bryson, A., Davis, J. C., & Freeman, R. (2016). It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States. *Journal of Labor Economics*, *34*(S2), S67–S97.

Belfield, C., Blundell, R., Cribb, J., Hood, A., & Joyce, R. (2017). Two decades of income inequality in Britain: the role of wages, household earnings and redistribution. Working paper, Institute for Fiscal Studies.

Benguria, F. (2015). Inequality Between and Within Firms: Evidence from Brazil. Working paper.

Blundell, R., Crawford, C., & Jin, W. (2014). What can wages and employment tell us about the UK's productivity puzzle? *Economic Journal*, *124*, 377–407.

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2016). Firms and labor market inequality: Evidence and some theory. NBER Working Paper 22850, National Bureau of Economic Research.

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics*, *128*(3), 967–1015.

Card, D., Lemieux, T., & Riddell, W. C. (2004). Unions and Wage Inequality. *Journal of Labor Research*, *25*(4), 519–562.

Cortes, G. M., & Salvatori, A. (2016). Delving into the Demand Side: Changes in Workplace Specialization and Job Polarization. Working paper.

Davis, S. J., & Haltiwanger, J. (1991). Wage Dispersion between and within U.S. Manufacturing Plants, 1963-86. *Brookings Papers on Economic Activity. Microeconomics*, *1991*, 115–200.

Devereux, P., & Hart, R. (2006). Real wage cyclicality of job stayers, within-company job movers, and between-company job movers. *Industrial and Labor Relations Review*, *60*(1), 105–119.

Dunne, T., Foster, L., Haltiwanger, J., & Troske, K. R. (2004). Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment. *Journal of Labor Economics*, *22*(2), 397–429.

Elsby, M., Shin, D., & Solon, G. (2016). Wage Adjustment in the Great Recession and Other Downturns: Evidence from the United States and Great Britain. *Journal of Labor Economics*, *34*(S1), S249 – S291.

Faggio, G., Salvanes, K. G., & Reenen, J. V. (2010). The evolution of inequality in productivity and wages: Panel data evidence. *Industrial and Corporate Change*, *19*(6), 1919–1951.

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *Review of Economics and Statistics*, *89*(1), 118–133.

Helpman, E., Itskhoki, O., Muendler, M.-A., & Redding, S. J. (2017). Trade and Inequality: From Theory to Estimation. *The Review of Economic Studies*, *84*(1), 357–405.

Hills, J., Brewer, M., Jenkins, S. P., Lister, R., Lupton, R., Machin, S., Mills, C., Modood, T., Rees, T., & Riddell, S. (2010). An anatomy of economic inequality in the UK: Report of the National Equality Panel. LSE Research Online Documents on Economics 28344, London School of Economics and Political Science, LSE Library.

Juhn, C., Murphy, K., & Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, *101*(3), 410–42.

Katz, L., & Murphy, K. M. (1992). Changes in Relative Wages, 1963 - 1987: Supply and Demand Factors. *Quarterly Journal of Economics*, *107*(1), 35–78.

Machin, S. (2011). Changes in UK Wage Inequality Over the Last Forty Years. In P. Gregg, & J. Wadsworth (Eds.) *The Labour Market in Winter*. Oxford University Press.

Machin, S., & van Reenen, J. (1998). Technology and Changes in Skill Structure: Evidence from Seven OECD Countries. *Quarterly Journal of Economics*, *113*(4), 1215–1244.

Mueller, H. M., Ouimet, P. P., & Simintzi, E. (2016). Within-Firm Pay Inequality. Working paper.

Nickell, S., & Quintini, G. (2003). Nominal Wage Rigidity and the Rate of Inflation. *Economic Journal*, *113*(490), 762–781.

Nordström Skans, O., Per-Anders, E., & Holmlund, B. (2009). Wage Dispersion Between and Within Plants: Sweden 1985-2000. In E. P. Lazaer, & K. L. Shaw (Eds.) *The Structure of Wages: An International Comparison*, (pp. 217–260). University of Chicago Press.

ONS (2016a). Annual Survey of Hours and Earnings, 1997-2015: Secure Access. [data collection]. 8th Edition SN: 6689, Office for National Statistics, UK Data Service.

ONS (2016b). New Earnings Survey Panel Dataset, 1975-2015: Secure Access. [data collection]. 6th Edition SN: 6706, Office for National Statistics, UK Data Service.

Piketty, T. (2013). *Capital in the Twenty-first Century*. Harvard University Press.

Song, J., Price, D. J., Guvenen, F., Bloom, N., & von Wachter, T. (2016). Firming Up Inequality. Working paper.

Weber-Handwerker, E., & Spletzer, J. R. (2016). The Role of Establishments and the Concentration of Occupations in Wage Inequality. In L. Cappellari, S. W. Polachek, & K. Tatsiramos (Eds.) *Inequality: Causes and Consequences (Research in Labor Economics)*, vol. 43, (pp. 167–193). Emerald Group Publishing Limited.

Appendix A. Further description of the data and sample construction

In what follows we give some brief additional details regarding the datasets used, and how we have constructed the sub-samples thereof. All the relevant documentation and variable descriptions attached to these datasets are publicly available from the UK Data Service. The ONS has also published various documents concerning the data quality and consistency of the NESPD and ASHE. We will publish our replication files for the analysis and sample construction.

We focus on methodological details through the period 1996-2015. From 1975 to 2003, under its guise as the New Earnings Survey, very little changed in the methodology and construction of the longitudinal panel dataset. Throughout this period, it should be a true random sample of all employees in employment, irrespective of employment status, occupation, size of employer etc. Given the legal obligation of employers to respond, and their use of payrolls, it has a very high response rate and is believed to be accurate. There is also no cumulative attrition from the panel, as any individual not included in the NESPD in any year, for whatever reason, remains in the sampling frame the following year. Conditional on a hundred percent response, the NESPD is a true one percent random sample of employees. However, there are two major sources of under-sampling, both occurring if individuals do not have a current tax record. This could occur for some individuals who have very recently moved job, or for those who earn very little (mostly part-time), and so do not have to pay tax or National Insurance. From 2004, the ASHE replaced the NESPD. This aimed to sample some of those employees under-represented in the NESPD. It added supplementary responses for those without a PAYE reference, and also attempted to represent employees whose jobs changed between the determination of the sampling frame in January and the reference period in April. Since the ONS states that the bias these amendments were introduced to address were actually small, we do not believe they could affect our results substantially. The ASHE also introduced some imputations, using similar matched 'donor' observations where responses were, for example, missing an entry of basic hours but had recorded pay. These imputations were added for weighting purposes, but throughout our analysis we ignore the weights in the ASHE. From 2005, a new questionnaire was also created which was intended to reduce the latitude for respondents' own interpretations of what was being asked of them. From 2007 there were further notable changes. Beforehand, occupations were classified as follows: either, if the respondent stated an employee's job had not changed in the past year, the previous year's occupational classification was applied. Otherwise, it was manually coded. Afterwards an automatic coding, text recognition, tool was used. "The effect of using ACTR was to code more jobs into higher paying occupations. The jobs that tended to be recoded into these higher paying occupations generally had lower levels of pay than the jobs already coded to those occupations. Conversely, they tended to have higher levels of pay than the other jobs in the occupations that they were recoded out of. The impact of this was to lower the average pay of both the occupation group that they had moved from and that they had moved to." As such, this would certainly increase within occupation wage inequality for the highest earners, and reduce it for the lowest earners. Nonetheless, we do not believe this is significant in affecting our results. In the main text, we focus the graphical analysis on changes across the period 1997-2007, but also find our results are unchanged for the periods 1996-2006 and 1996-2001. From 2007, the sample size of the ASHE was reduced by twenty percent, with reductions targeted on those industries that exhibit the least variation in their earnings patterns. However, we do not believe this could have affected our results substantially.

To construct the sub-samples from the panel dataset for 1975-2015, for the analysis of hourly or weekly pay, we first drop a few cases of duplicates over all variables. Then, using the panel identifier, year, the information from the IDBR concerning enterprise status and number of employees, industry classification and gross weekly pay including overtime, we also drop some cases which are then determined to be the same job. We do not drop observations where an individual has multiple jobs. We keep only observations for individuals aged 16-64, and which have not been marked as having loss of pay in the reference period through absence, employment starting in the period, or short-time working, and which are marked as being on an adult rate of pay (i.e. dropping trainees and apprenticeships). This is practically the same filter applied for ONS published results using the NESPD or ASHE. We also drop all observations with zero or missing values for basic hours, and hourly or weekly pay excluding overtime. Basic hours are intended to be a record for the employee in a normal week, excluding overtime and meal breaks. Gross weekly pay is the main recorded value in the survey, and from this overtime records are then simply subtracted. Hourly rates are then derived from dividing by basic hours worked. We drop observations with over a hundred basic hours worked, as these could reflect measurement error and inclusion of overtime. Full-time is defined as working over thirty basic hours in a week. But there are a tiny number of discrepancies in some years, we believe relating to teaching contracts, where the definition applied by the ONS differs. We however recode these such that for all observations the thirty hours threshold applies. To further address some potential for measurement error especially in the recorded basic hours, we drop observations whose hourly rate of pay excluding overtime is less than eighty percent of the National Minimum Wage (NMW) which applies each April, with allowance for the different age-dependent rates of the NMW over time. We set the threshold lower to avoid dropping observations where employers have rounded figures about the NMW, where the degree of rounding could vary with the actual value of the NMW, a behaviour which has been hypothesised by the ONS. To then construct the large firm sample, we drop all employers whose exact enterprise reference number of employees from the IDBR, which is only available from 1996 onwards, is less than 250. We also drop observations where the IDBR status, number of employees or industry classification is missing. We then identify each employer in the dataset using the combination of their five digit industry code, IDBR status and exact number of IDBR enterprise employees, within each year. For large firms we are confident this can uniquely identify the reporting organisation of the NESPD. The large firm samples we subsequently analyse then condition on there being a minimum number of remaining job observations per firm in a year. For annual pay, we construct the large firm samples in the same way, except we additionally filter out observations where the employee is reported to not have been with the employer for twelve months, and drop observations with zero or missing values of annual gross pay in place of hours or weekly pay. When handling the ASHE annual cross-section datasets we use the exact same approach, except here there is a unique enterprise level identifier which we can use to identify the firms within each year.

For 1996-2001 occupations are classified using the three digit ONS 1990 Standard Occupational Classification (SOC). For 2002-2010, this is replaced with the four digit SOC 2000, and for 2011-2015, with the SOC 2010. We experimented using the ONS' publicly available cross-walk from 2010 and 2000 to 1990 classification, but discovered that this causes a large structural break in the distribution of occupations. In particular, it causes a substantial additional degree of polarisation of work from 2002 onwards, which would potentially generate erroneous and large increases in within occupation inequality around this date. To address this we rely on a conversion of SOC 1990 and 2000 to the 1988 International Standard Classification of Occupations (ISCO). We obtain these conversions from the Cambridge Social

Interaction and Stratification Scale (CAMSIS) project. For the industry classification, we convert ONS Standard Industrial Classification (SIC) 2007 to 2003, using files made available by the UK Data Service. This conversion uses the 2008 Annual Respondents Dataset where both classifications were applied, and where any 2007 code mapping to multiple 2003 codes is decided using whichever of the two bore a greater share of economic output. For 1996-2002, the work region of the employee is missing, and so we derive this ourselves consistent with the ONS geo-maps, using the more detailed work area variable.

TABLE A1: Comparison of baseline sample firm size distribution, and represented employees, with UK population of enterprises, 2013

| Enterprise size[†] | Number of obs. | | Total employees in enterprises (000s)[†] | |
	Sample firms[‡]	UK enterprises	Sample firms	UK enterprises
250 - 999	92	6,400	69	2,927
1,000 - 1,999	308	1,050	469	1,455
2,000 - 4,999	644	830	2,098	2,612
5,000+	596	635	8,204	8,805
Total	1,640	8,915	10,804	15,799

[†] Values for sample firms use the IDBR record of the number of employees in the enterprise which includes the firm. This is not the number of observations of employee jobs in the sample.
[‡] All firms in the baseline sample with a minimum of ten full-time employee observations in the NESPD in 2013, and subject to the other sampling criteria described in the text.
 Notes.- author calculations using NESPD. UK enterprises population figures from UK Business: Activity, Size and Location (IDBR, March 2015).

TABLE A2: Baseline sample number of firm and employee observations by employer size, 1997 & 2007

| Enterprise size[†] | Firms | | | Employees | | | IDBR ent. employees (000s)[†] | | |
	1997	2007	Change in share	1997	2007	Change in share	1997	2007	Change in share
250 - 999	125	43	-0.05	1,729	497	-0.03	89	32	0.00
1,000 - 1,999	352	214	-0.08	5,322	2,814	-0.06	539	328	-0.02
2,000 - 4,999	512	548	0.05	10,068	9,789	-0.03	1,612	1,817	0.03
5,000+	485	569	0.09	26,915	36,242	0.12	9,431	8,350	-0.01
Total	1,474	1,374		44,034	49,342		11,671	10,525	

[†] Values for the sample firms use the IDBR record of number of employees in the enterprise which includes the firm.
 Notes.- author calculations using NESPD.

TABLE A3: Baseline sample incidence of ISCO88 major occupation groups

Major group[†]	1997	2007	Change
1	0.12	0.15	-0.02
2	0.22	0.16	0.06
3	0.09	0.14	-0.05
4	0.24	0.20	0.03
5	0.11	0.14	-0.03
6 & 7	0.08	0.04	0.04
8	0.08	0.05	0.03
9	0.05	0.11	-0.06

[†] Key: 1. Legislators, senior officials and managers, 2. Professionals, 3. Technicians and associate professionals, 4. Clerks, 5. Service workers and shop and market sales workers, 6. Skilled agricultural and fishery workers, 7. Craft and related trades workers, 8. Plant and machine operators and assemblers, 9. Elementary occupations.
Notes.- author calculations using NESPD.

Appendix B. Mathematical details

B.1 Variance decomposition - hours and wages

From the main text, we can re-write (1), the total variance of log weekly wages as follows, where ω and η denote the hourly non-log wage rate and hours worked respectively, and h denotes log hours,

$$\underbrace{\frac{1}{N}\sum_j^J\sum_i^{N_j}\left[\ln(\omega_{ij}\eta_{ij})-\overline{\ln(\omega_{ij}\eta_{ij})}\right]^2}_{\text{Overall - }V_e} = \underbrace{\frac{1}{N}\sum_j^J\sum_i^{N_j}\left[\ln(\omega_{ij}\eta_{ij})-\overline{\ln(\omega_j\eta_j)}\right]^2}_{\text{Within-firm - }V_{wf}} + \underbrace{\sum_j^J\frac{N_j}{N}\left[\overline{\ln(\omega_j\eta_j)}-\overline{\ln(\omega\eta)}\right]^2}_{\text{Between-firm - }V_{bf}},$$

(5)

with

$$V_{wf} = \underbrace{\frac{1}{N}\sum_j^J\sum_i^{N_j}\left[w_{ij}-\overline{w_j}\right]^2}_{V_{wf}^w} + \underbrace{\frac{1}{N}\sum_j^J\sum_i^{N_j}\left[h_{ij}-\overline{h_j}\right]^2}_{V_{wf}^h} + \underbrace{\frac{2}{N^2}\sum_j^J\sum_i^{N_j}\left[\left(w_{ij}-\overline{w_j}\right)\left(h_{ij}-\overline{h_j}\right)\right]}_{2Cov_{wf}(w,h)} \quad (6)$$

and

$$V_{bf} = \underbrace{\sum_j^J\frac{N_j}{N}\left[\overline{w_j}-\overline{w}\right]^2}_{V_{bf}^w} + \underbrace{\sum_j^J\frac{N_j}{N}\left[\overline{h_j}-\overline{h}\right]^2}_{V_{bf}^h} + \underbrace{2\sum_j^J\frac{N_j}{N^2}\left[\left(\overline{w_j}-\overline{w}\right)\left(\overline{h_j}-\overline{h}\right)\right]}_{2Cov_{bf}(w,h)}. \quad (7)$$

B.2 Shift-share analysis of the change in the firm component of employee wages

Let each decile be denoted by d, where N^d is all employees observed in a period in that decile of the wage distribution. Let k denote an employment type, with K types in total. The share of all employees, irrespective of decile, in type k in the firm of an employee i is given by $\alpha_{k,i}$. The mean log wage of type k in the firm of employee i is given by $w_{k,i}$. We let this value be zero where a firm does not employ anybody of type k. We can write the mean of firm average log wages for employees in a decile as

$$
\frac{1}{N^d} \sum_{i \in d}^{N^d} \{\bar{w}_j\}_i = \frac{1}{N^d} \sum_{k}^{K} \sum_{i \in d}^{N^d} \alpha_{k,i} w_{k,i}
$$

$$
= \sum_{k}^{K} \left[\underbrace{\left(\frac{1}{N^d} \sum_{i \in d}^{N^d} \alpha_{k,i} \right)}_{\bar{\alpha}_k} \underbrace{\left(\frac{1}{N^d} \sum_{i \in d}^{N^d} w_{k,i} \right)}_{\bar{w}_k} + \underbrace{\frac{1}{N^d} \sum_{i \in d}^{N^d} (\alpha_{k,i} - \bar{\alpha}_k)(w_{k,i} - \bar{w}_k)}_{Cov(\alpha_k, w_k)} \right]. \tag{8}
$$

Using (8) and denoting historical values by $'$, we can write the difference in the mean of firm average log wages for employees in some decile, between period t and some historical period, representing the difference operator by Δ, as

$$
\sum_{k}^{K} \left[\underbrace{\bar{\alpha}'_k \Delta \bar{w}_{k,t}}_{\text{Wages effect}} + \underbrace{\bar{w}'_k \Delta \bar{\alpha}_{k,t}}_{\text{Shares effect}} + \underbrace{\Delta \bar{\alpha}_{k,t} \Delta \bar{w}_{k,t}}_{\text{Interaction effect}} + \underbrace{\Delta Cov(\alpha_{k,t}, w_{k,t})}_{\text{Covar. effect}} \right]. \tag{9}
$$

31

Appendix C. Additional figures

FIGURE C1: Shares of firms and employees in the baseline sample in SIC 2003 sectors, 1997 & 2007

(a) Firms

(b) Employees
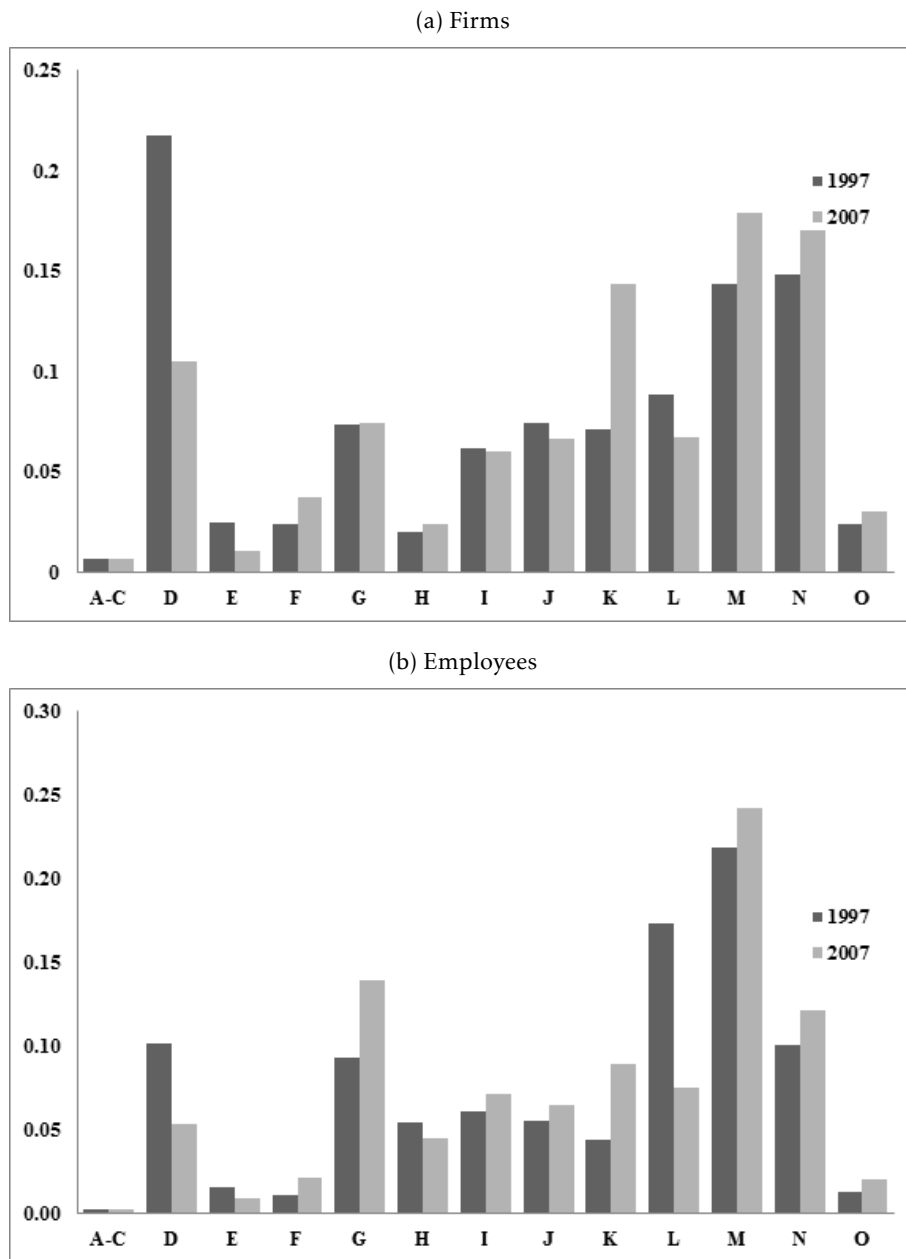
Notes.- author calculations using New Earnings Survey. See the text for description of the sample. SIC2003 codes: A. Agriculture, hunting and forestry; B. Fishing, C. Mining and quarrying, D. Manufacturing, E. Utilities, F. Construction, G. Wholesale and retail, H. Hotels and restaurant, I. Transport and telecommunication, J. Financial intermediation, K. Real estate, business services, L. Public admin and defence, social security, M. Education, N. Health and social work, O. Other comunnity and social services.

FIGURE C2: Mean of real log wages in large firms, full-time employees only, and comparison the with whole NESPD sample, 1975-2015

(a) Weekly

(b) Hourly

(c) Annual
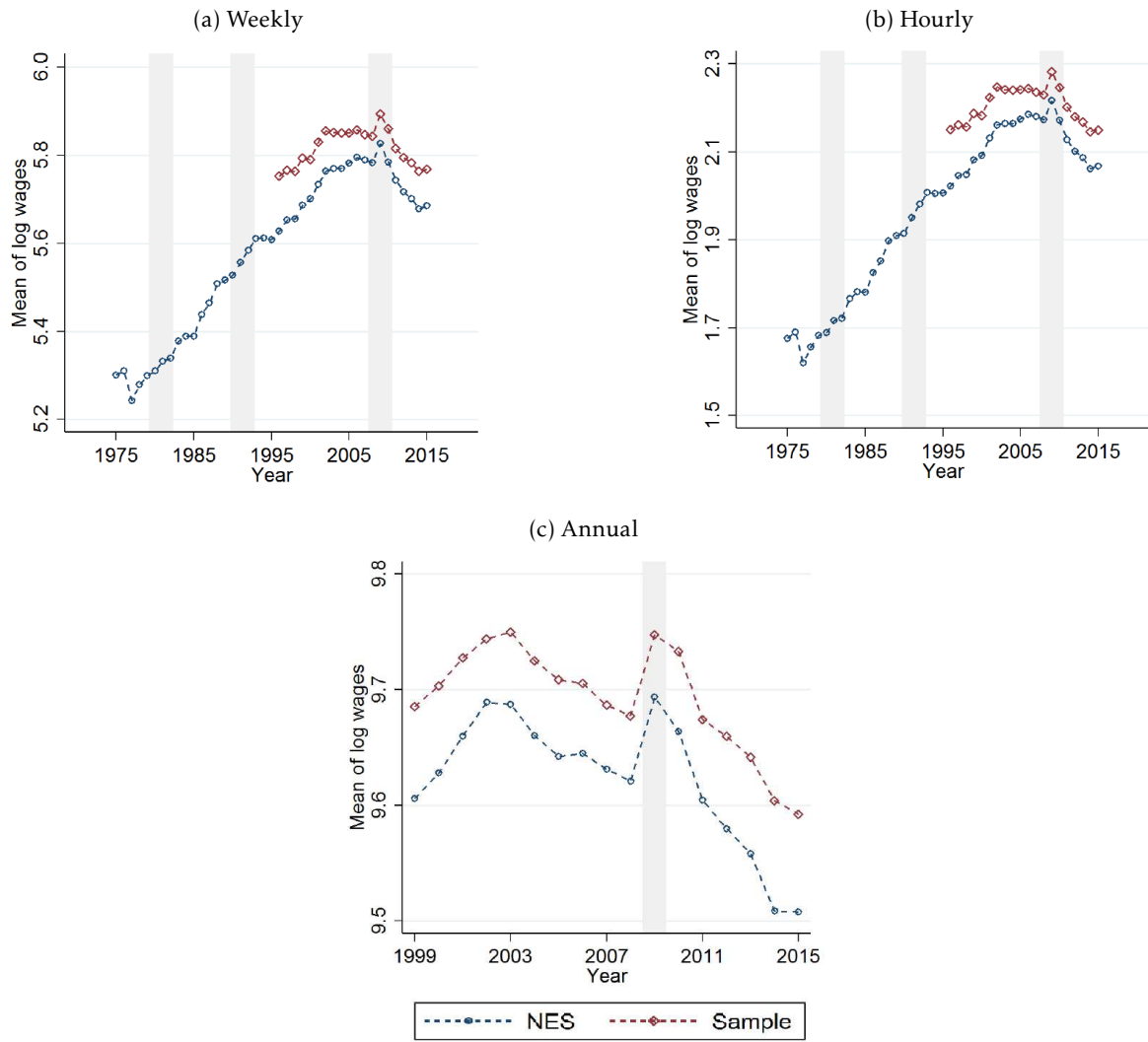
Notes.- see Figure 1. The top one percent of wage observations in any year are excluded from all calculations here.

FIGURE C3: Mean of real log wages in large firms, all employees, and comparison with the whole NESPD sample, 1975-2015

(a) Weekly

(b) Hourly

(c) Annual
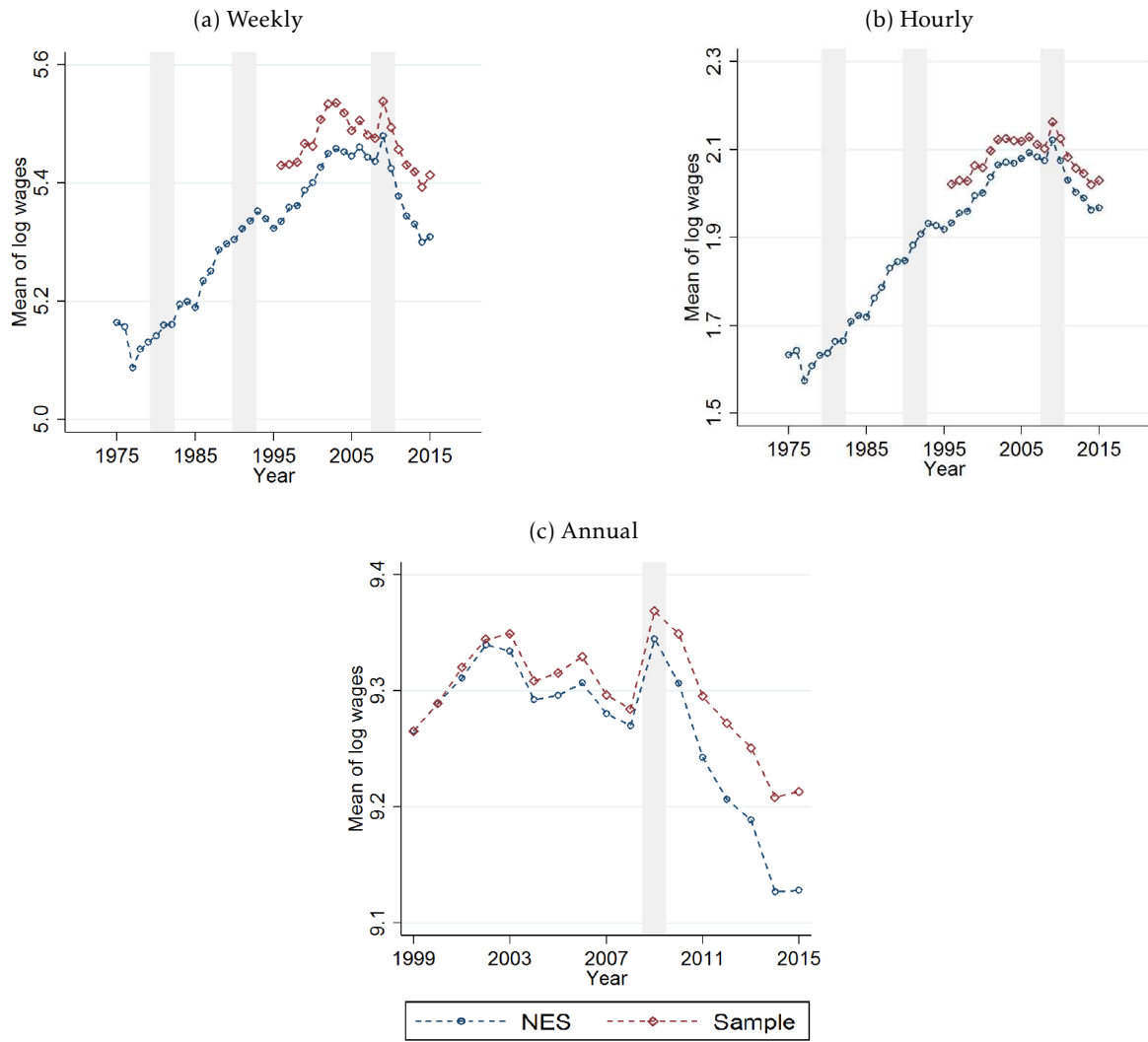
Notes.- see Figure 1, except here is with all employees. The top one percent of wage observations in any year are excluded from all calculations here.

FIGURE C4: Percentiles of real log wages in large firms, all employees, and comparison with the whole NESPD sample, 1975-2015

(a) Weekly

(b) Annual
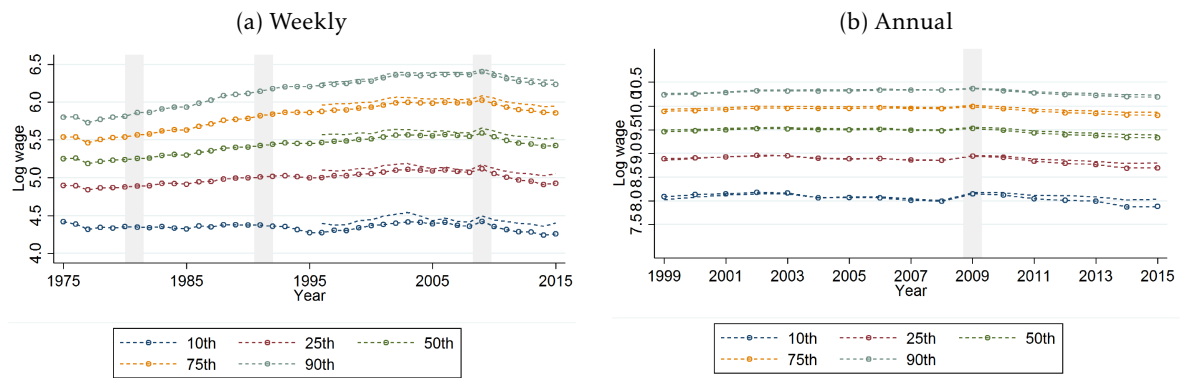
Notes.- see Figure 1, except here is with all employees. Dashed lines without markers are the series for the large firm sample of the NESPD.

FIGURE C5: Percentiles of real log wages in large firms, all employees: differences relative to 1996/9

(a) Weekly

(b) Annual
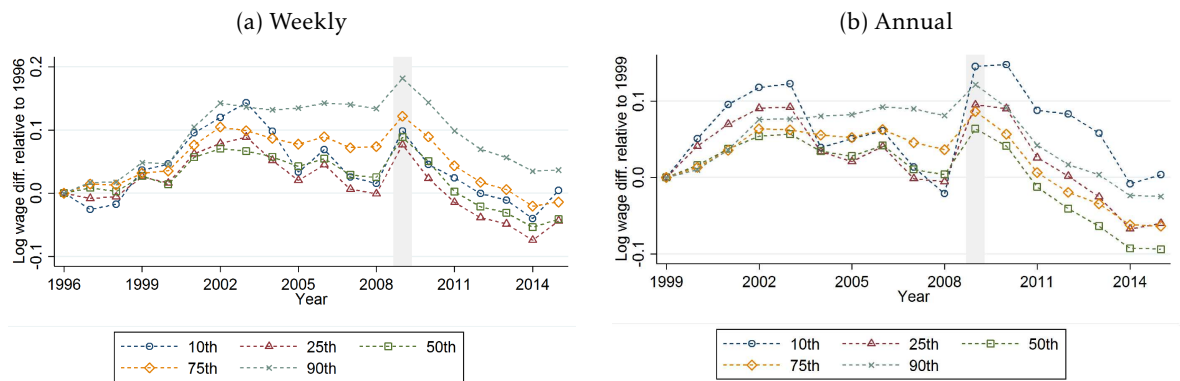
Notes.- see Figure 2, except here is with all employees.

FIGURE C6: Share of variance in log weekly employee wages from within-firm component, 1996-2015: comparison of firm weights

Notes.- see Figure 3. 'Sample' gives results where firms are weighted using their share of sample observations in that year. 'IDBR...' gives results where firms are weighted using their administrative record of enterprise size from the IDBR.

FIGURE C7: Share of variance in log weekly employee wages from within-firm component

(a) All firms

(b) Private sector excl. wholesale, retail, hotels, restaruants etc.

(c) Public sector only
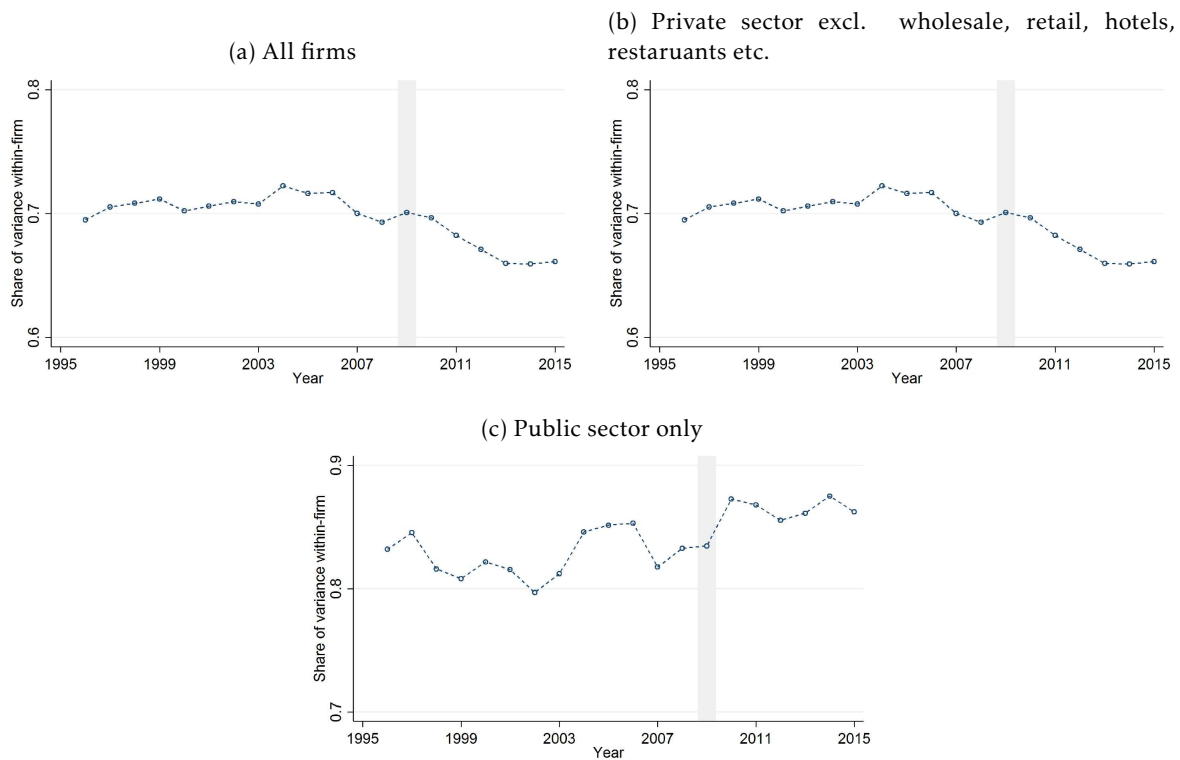
Notes.- see Figure 3. Panel (b) excludes major SIC 2003 sectors G & H. Public sector is represented by public corporation or nationalised industry, central government and local authority employers.

FIGURE C8: Share of variance in log hourly employee wages from within-firm component

(a) All firms

(b) Private sector excl. wholesale, retail, hotels, restaruants etc.

(c) Public sector only
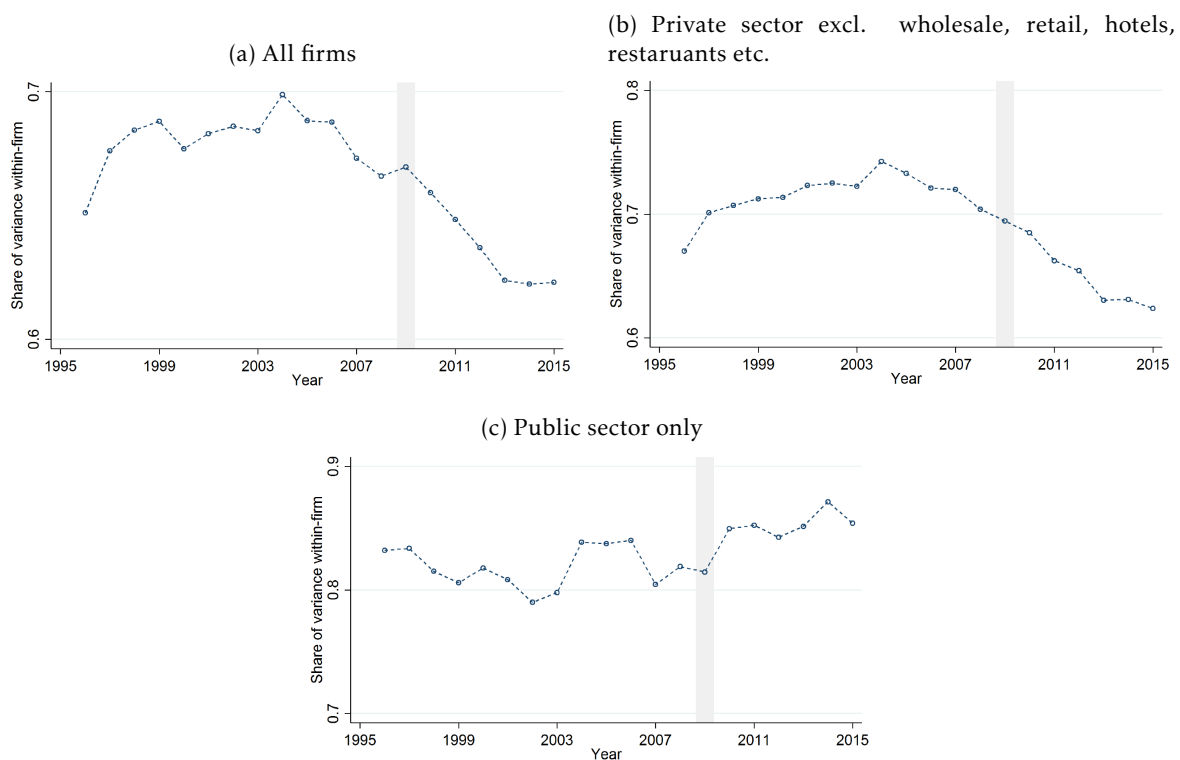
Notes.- see Figure C7

FIGURE C9: Share of variance in log annual employee wages from within-firm component

(a) All firms   (b) Public sector only

Notes.- see Figure 3 and see Figure C7.

FIGURE C10: Share of variance in log weekly employee wages from within-firm component: NESPD large firm sample vs ASHE enterprises
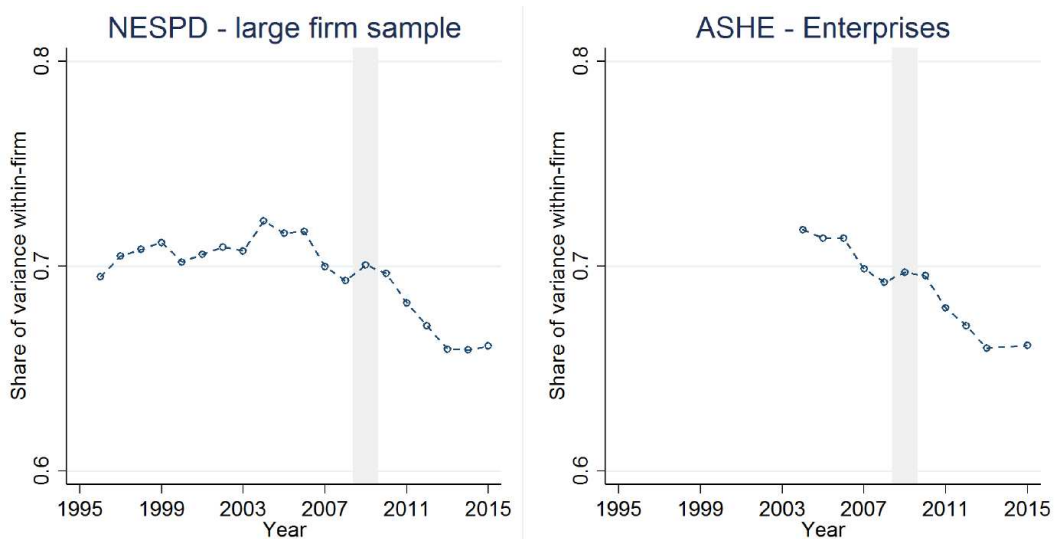
Notes.- author calculations using the New Earnings Survey and Annual Survey of Hours and Earnings, age 16-64 only, all employees. Weekly wages exclude overtime. In the left panel the data is for all large firms in the NESPD who have at least ten employee observations in a year. The right panel is the equivalent but using IDBR enterprise identifiers in the ASHE, instead of a broader definition of a 'firm'. The top one percent of wage values in each year are excluded from calculations here. Shaded areas represent official UK recessions.

37

FIGURE C11: Change 2008-2015 in the average log weekly wage by percentile of employees and the contribution from firms: NESPD large firm sample vs ASHE large enterprises
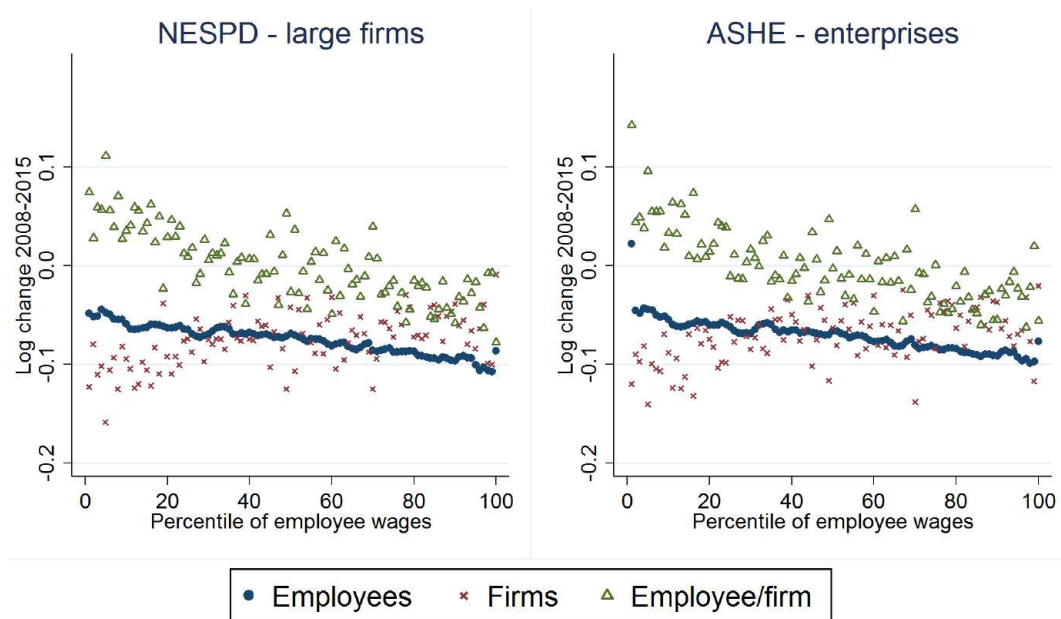
Notes.- see Figure 7 and Figure C10.

FIGURE C12: Change in the average residual log weekly wage by percentile of employees and the contribution from firms: other ten year time periods
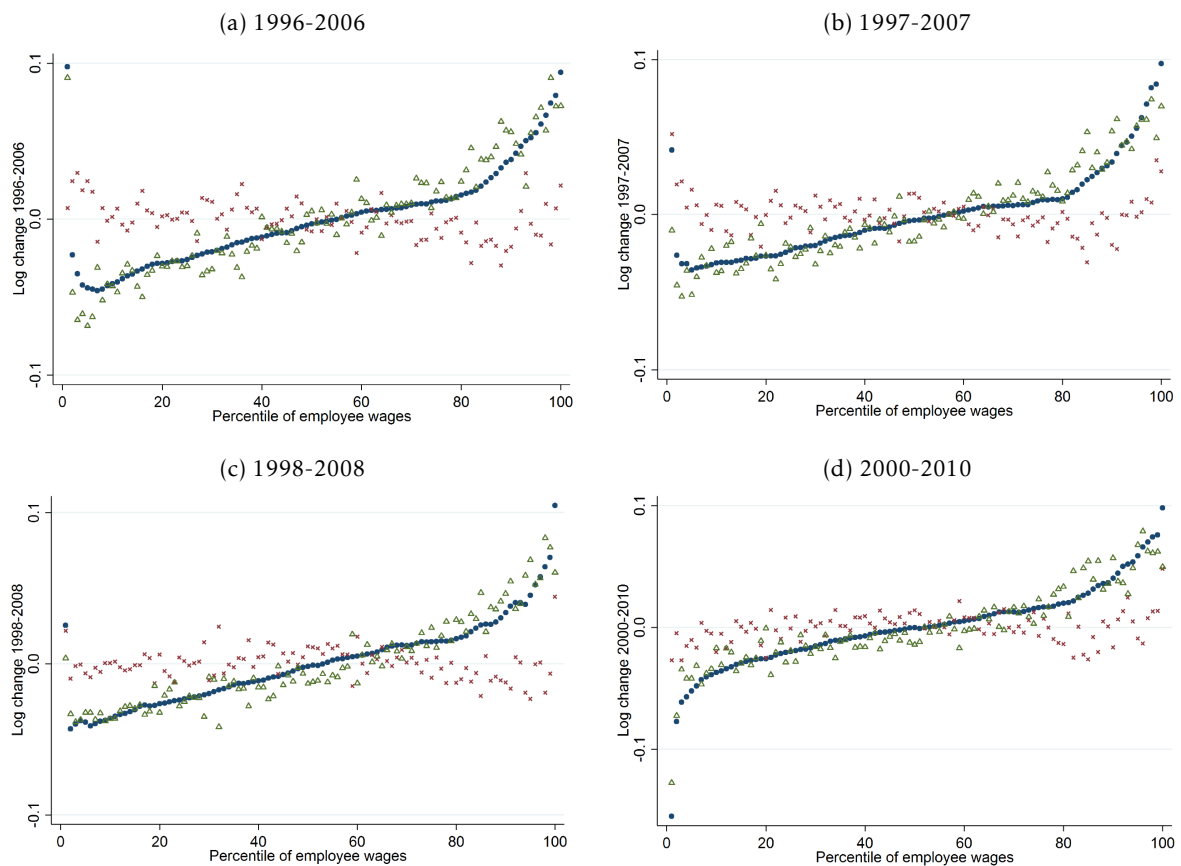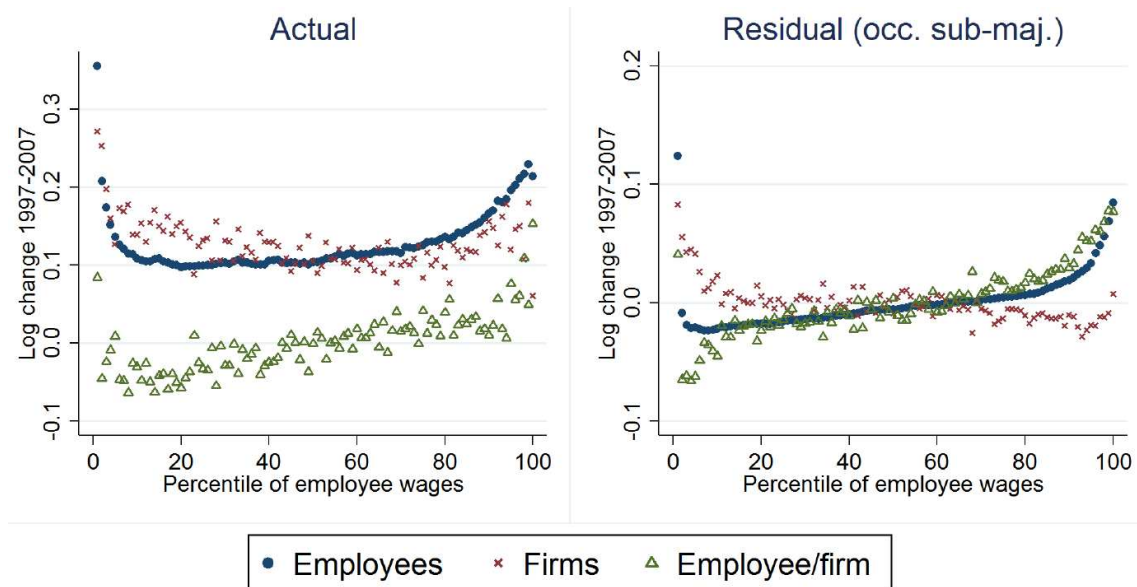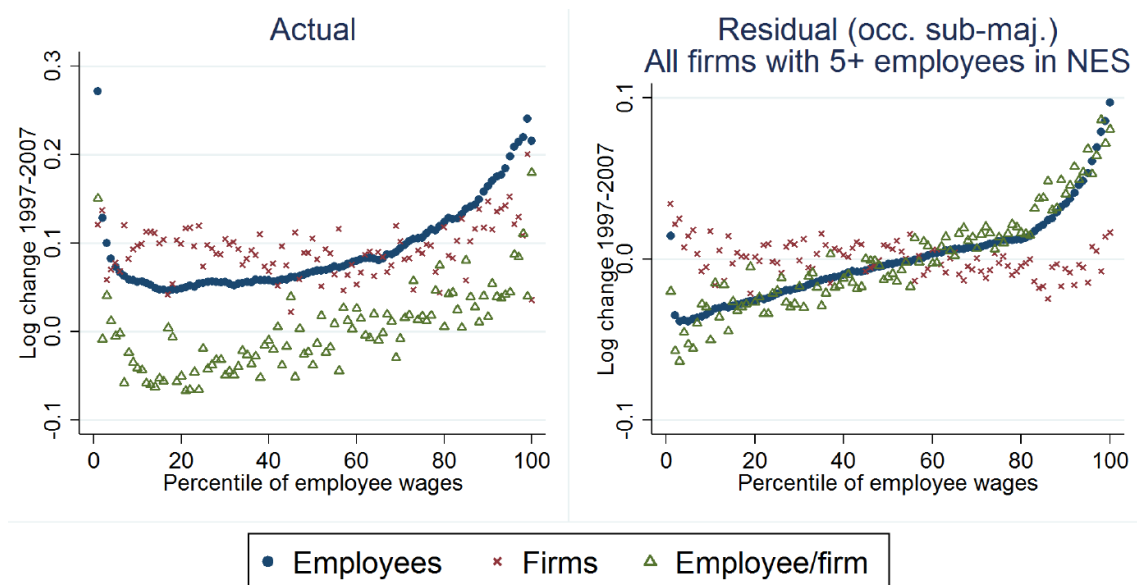

Notes.- see Figure 7 and Figure 8. Residual log wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

FIGURE C13: Change in the average residual log weekly wage by percentile of employees and the contribution from firms: other five year time periods
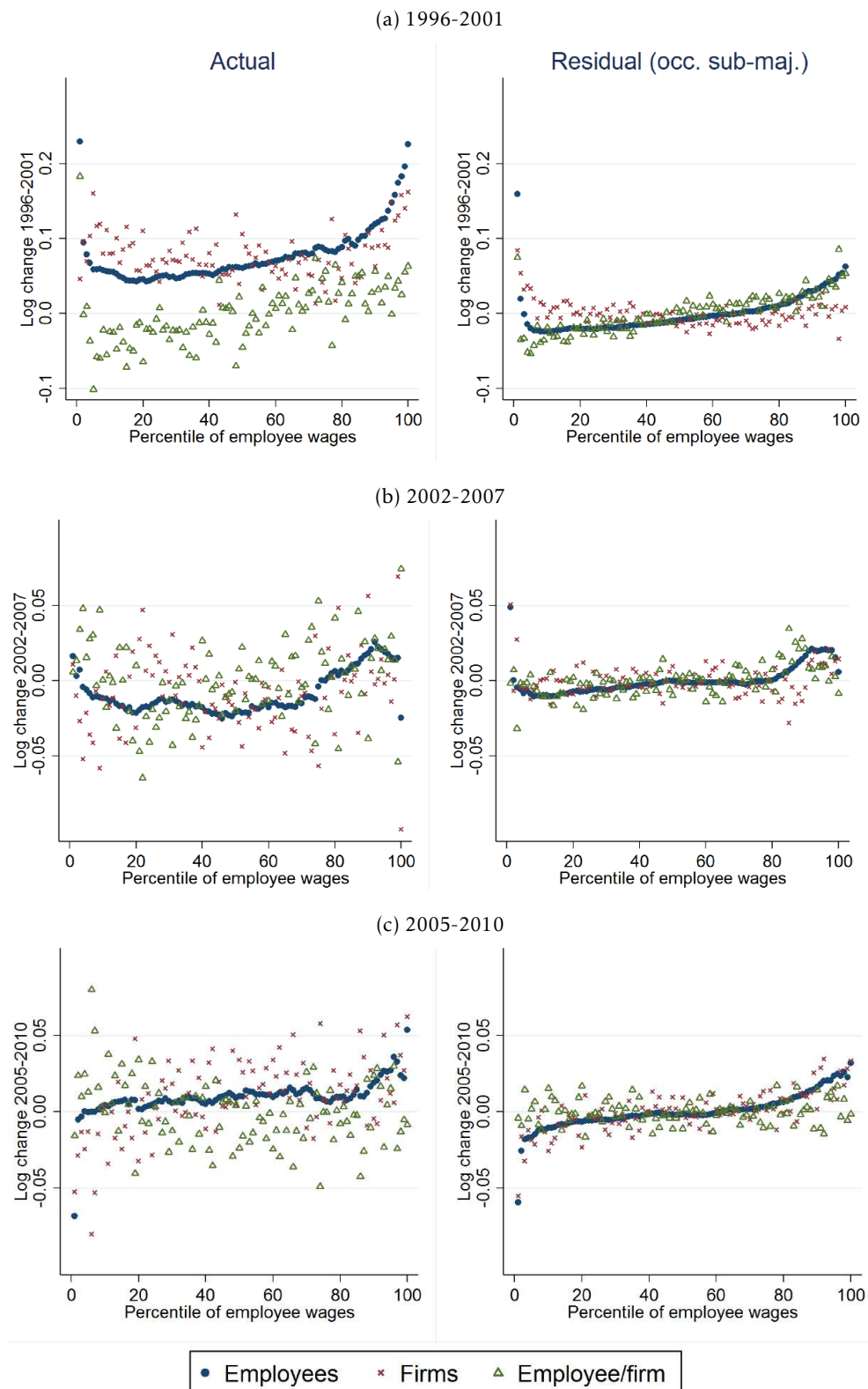
(a) 1996-2001

(b) 2002-2007

(c) 2005-2010

• Employees   × Firms   △ Employee/firm

Notes.- see Figure 7 and Figure 8. Residual log wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

39

FIGURE C14: Change 1997-2007 in the average residual log weekly wage by percentile of employees and the contribution from firms: all large firms in the NESPD with 1+ employee observations

Notes.- see Figure 7 and Figure 8. The data used here is for all large firms who have at least one employee observation in the NESPD in a year. Residual log wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

FIGURE C15: Change 1997-2007 in the average residual log weekly wage by percentile of employees and the contribution from firms: all large firms in the NESPD with 5+ employee observations
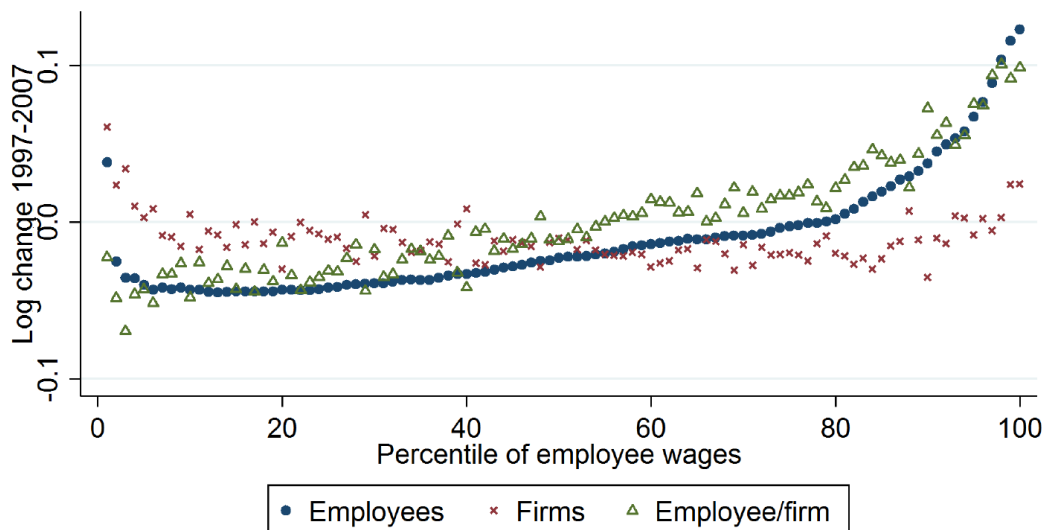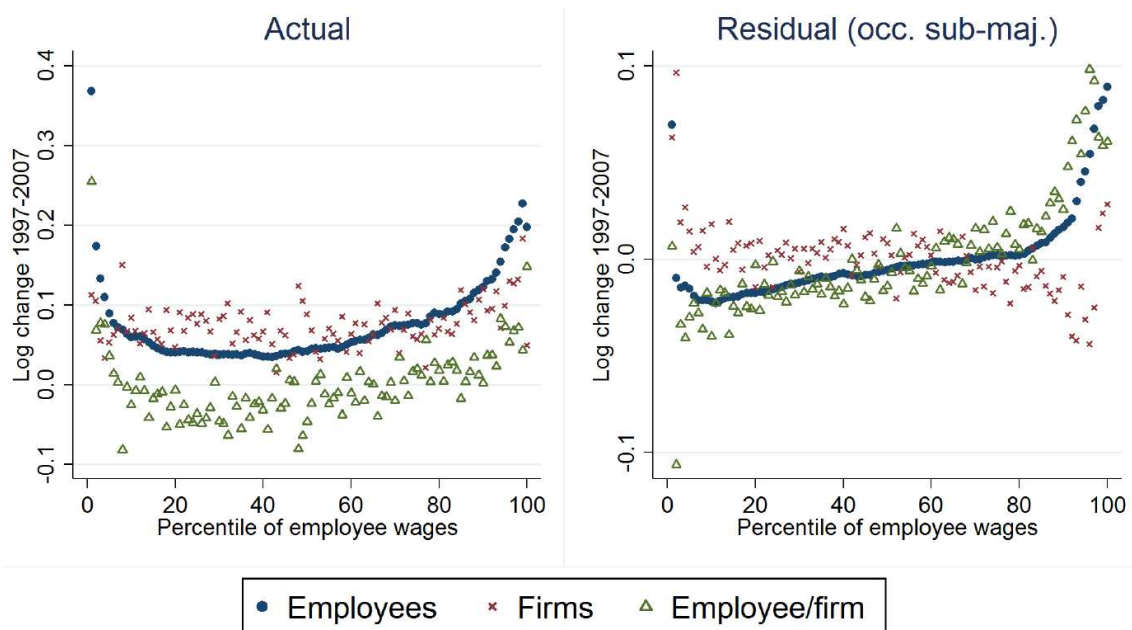

Notes.- see Figure C14, except the data used here is for all large firms who have at least five employee observations in the NESPD in a year.

FIGURE C16: Change 1997-2007 in the average residual log weekly wage by percentile of employees and the contribution from firms: all large firms in the NESPD with 20+ employee observations
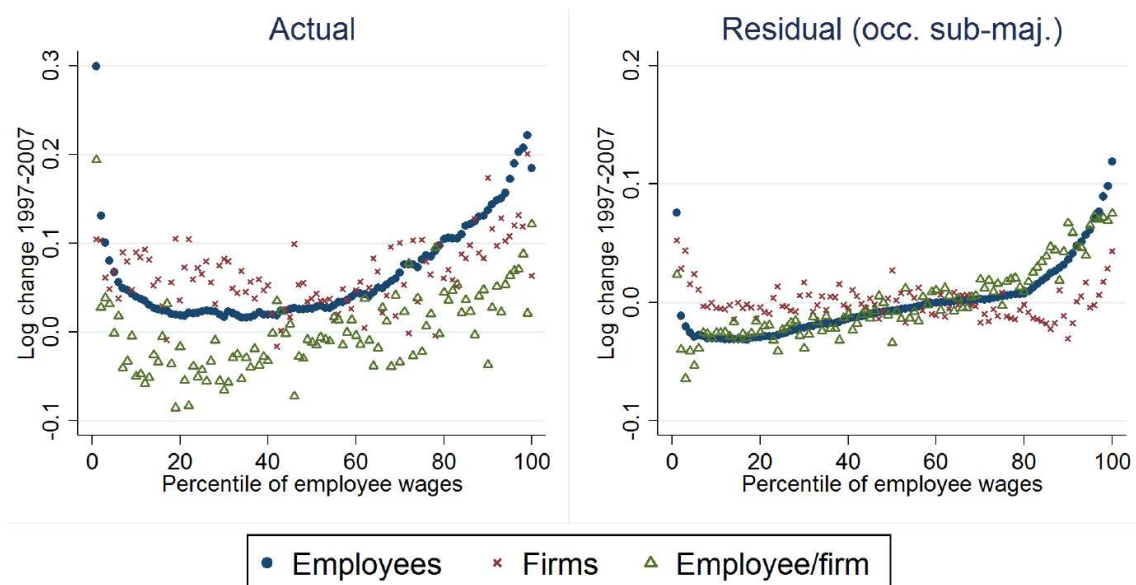
Notes.- see Figure C14, except the data used here is for all large firms who have at least twenty employee observations in the NESPD in a year.

FIGURE C17: Change 1997-2007 in the average residual log weekly wage by percentile of employees and the contribution from firms: private sector only
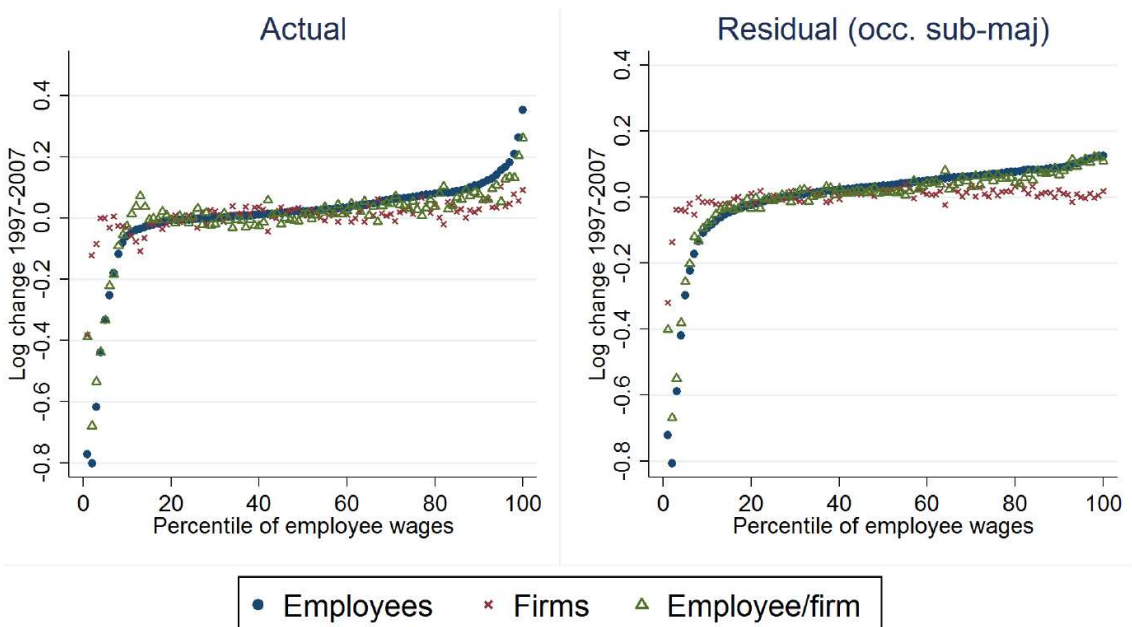

Notes.- see Figure 7 and Figure 8, except the data used here is for all large private sector firms in the NESPD who have at least ten full-time employee observations in a year. Residual log wages are estimated using OLS for all firms, with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

FIGURE C18: Change 1997-2007 in the average log hourly wage by percentile of employees and the contribution from firms: comparison with residual wages
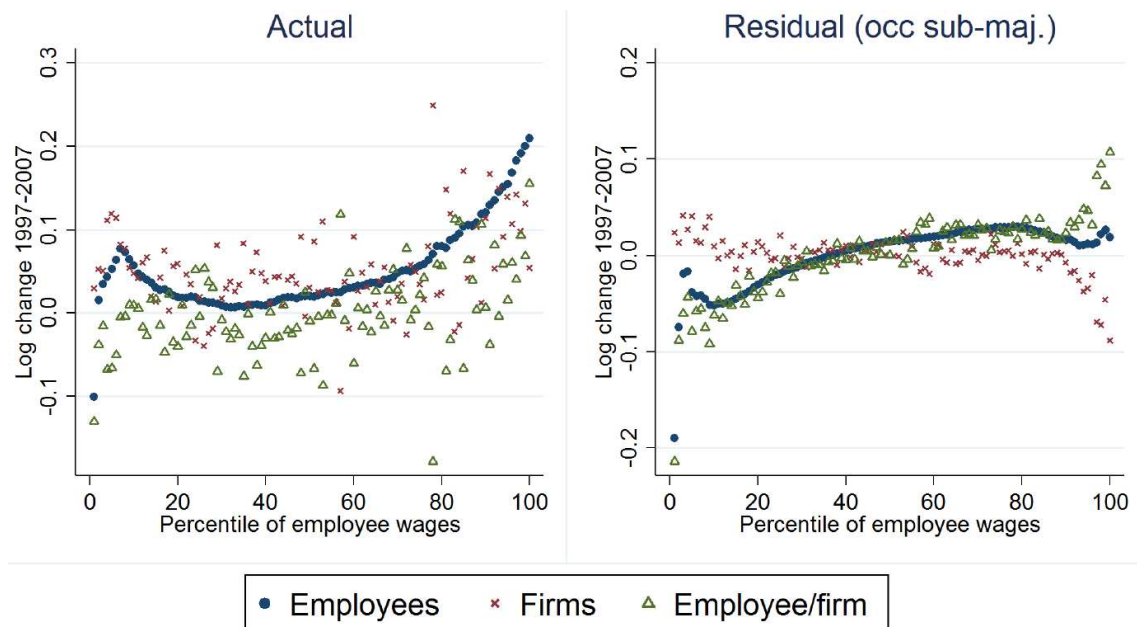

Notes.- see Figure 7 and Figure 8. Hourly wages exclude overtime. Residual log hourly wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

FIGURE C19: Change 1997-2007 in the average log annual wage by percentile of employees and the contribution from firms: comparison with residual wages

Notes.- see Figure 7 and Figure 8, except the data used here is for all large firms who have at least ten employee observations, who have been with the firm at least a year, in the NESPD in a year. Residual log wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).

FIGURE C20: Change 1997-2007 in the average log weekly wage by percentile of employees and the contribution from firms, full & part-time workers: comparison with residual wages

Notes.- see Figure 7 and Figure 8, except here the data is for all employees, not full-time only. Residual log wages are estimated using OLS with controls for sex, age, age squared, major regions and occupation sub-major groups (ISCO88).